EPH - International Journal of Educational Research
ISSN (Online): 2208-2204
Volume 01 Issue 02 May 2017

DOI: https://doi.org/10.53555/ephijer.v1i1.1

MULTILEVEL MODELLING OF GENERAL CERTIFICATE OF EDUCATION ORDINARY LEVEL PERFORMANCE OF SCHOOLS IN CONFLICT AFFECTED AREAS IN SRI LANKA

D.G.I. Kulawardana[1*], M.R. Sooriyarachchi[2]
*[*1,2]Department of Statistics, Faculty of Science, University of Colombo, Colombo 03, Sri Lanka*
[2]Email: roshinis@hotmail.com

Corresponding Author: -
Email: dgik_8@yahoo.com

Abstract: -
Multilevel data structures are known as consisting of multiple units of analysis, one clustered within the other. The concept of multilevel data modelling has been developed for several years mainly because the researchers have realized the disadvantages of ignoring such multilevel data structures. This study aims to find out factors affecting the General Certificate of Education Ordinary Level (G.C.E. O/L) pass rate at schools located in civil war affected provinces in Sri Lanka. The study also extends to observe the multilevel data structure by schools, districts and provinces, and determine how these levels have an impact on the G.C.E. O/L pass rate. The above has been undertaken by the application of advanced analysis focused on developing Generalized Linear Multilevel Model for ordered categorical response using the Bayesian Markov Chain Monte Carlo estimation method employing MLwiN 2.10 software. Finally, the partial non-proportional odds model was selected as the most appropriate model for the Educational data used in this study based on account of its simplicity and accuracy.

Keywords:- *Bayesian Method, Education, Multilevel data modelling, Ordinal categorical response, Partial non-proportional odds model*

1. INTRODUCTION

Multilevel data structures are commonly observed phenomena in many fields. The importance of multilevel data structures had been well identified prior to application of powerful analytical tools for multilevel modelling. Most developments of this type of modelling have been focused on a continuous response variable rather than categorical response variables. Hence, Multilevel modelling for an ordered categorical response is somewhat of a novel application [1]. Also, the interest in modelling for ordinal categorical response is often encountered in Educational, Social and Medical data [2]. The advanced analysis of this study is based on multilevel modelling for ordered categorical response with Educational data. And the focused area of this study is on G.C.E. O/L performance of schools in conflict affected areas, namely, the Northern and Eastern provinces in Sri Lanka during the war period and the post-war period.

Sri Lanka experienced a devastating armed civil conflict over a period of 30 years ending in May 2009. The Northern and Eastern Provinces were the most affected by this civil war. In times of conflict, children were adversely affected through them being recruited as child soldiers, disrupting of their education and losing their homes and other properties. After the war, the Government of Sri Lanka (GOSL) as well as different Organizations of the UN and NonGovernment Organizations implemented diverse projects for economic and social uplift of the Northern and Eastern Provinces. The GOSL launched two large projects named *Uthuru Wasanthaya* and *Negenahira Navodaya* aimed at resettlement of people and rehabilitation of economic infrastructure to re-establish people's livelihoods and provide basic needs such as transport, water, sanitation, health and education in the above two Provinces [3].

Moreover, Education is an essential factor in the socio-economic transformation of a country. Also, it is generally expected that the exposure to armed civil conflict mostly affects the educational attainment of affected children [4]. As many researchers have observed, it is difficult to alleviate poverty without imparting education at least up to the G.C.E. O/L standard [5].

1.1 Literature Review

1.1.1 Multilevel Modelling: Revisited

Subsequent to researchers engaging in developing systematic approaches for statistical modelling and analysis of hierarchically structured data, the concept of hierarchical data modelling commenced in the mid-1980s [6]. Accordingly, there was a consistent development for hierarchical data modelling as a result of the early work of Aitkin et al. (1986) on teaching styles data and Aitkin's continuous classic work with Longford (1986) [7,8]. By the early 1990s, an important set of techniques, experiences and software packages that could be applied consistently, had been developed.

1.1.2 Importance of Multilevel Modelling

Researchers have developed elaborate procedures for statistical multilevel modelling considering different types of response variables. Mainly, analysis of multilevel modelling has several benefits. Firstly, it helps to provide statistically efficient estimates of regression coefficients and also using clustering information, it provides accurate standard errors, confidence intervals as well as significant tests. Secondly, it allows the use of covariates at any level of a hierarchy. An example for this fact is to help the researcher to discover the extent to which differences in average of students' performance between schools is accounted for by factors such as characteristics of students. Also, it helps the researcher to find out whether the variation between schools is greater for students who have initially high marks than for students who have initially low marks. Finally, it enables to rank individual schools by considering performance of students after adjusting for intake achievements [9].

1.1.3 MLwiN Software used for Multilevel Models

Multilevel modelling had been initially focused on a continuous variables. Further, multilevel theory and its implementation in software for binary, nominal and ordinal data were developed in the early 1990s. The outcome of this was the MLwiN software package that has been developed as the most sophisticated software to accommodate the new applications of multilevel modelling [1]. In this software package, there are some advanced features which are not seen in other statistical packages. Especially, MLwiN facilitates Markov Chain Monte Carlo (MCMC) estimation for a range of statistical models. It provides many user friendly facilities when fitting multilevel models. Further, MLwiN 2.10 version contains enhanced features like estimation, exploring, importing and exporting data as well as with other improvements for ease of use [10].

In this study, therefore, MLwiN 2.10 has been employed to fit multilevel models.

1.2 Objective

The objectives of the present study are to find out factors affecting G.C.E. O/L pass rate of Schools in Northern and Eastern Provinces of Sri Lanka while observing multilevel data structure into Schools, Districts and Provinces levels, and determining how these levels impact on the G.C.E. O/L pass rate. Accordingly, it is attempted to observe whether there is a significant improvement of the G.C.E. O/L pass rate of schools during the post-war period compared with the war period in the Northern and Eastern Provinces of Sri Lanka, and how the G.C.E. O/L pass rate of schools varies among districts in the Northern and Eastern Provinces.

2. Materials & Method

2.1 Data

The study is primarily based on the secondary data from two sources; the Department of Examinations and the Data Management Branch of the Ministry of Education. The former provided the G.C.E. O/L pass rates of schools and the

latter furnished the Annual School Census data. Data was collected based on accessibility from these two data sources. The data set consisted of G.C.E. O/L pass rates of schools from 2005 to 2014 by district, province, year (from the former source) and characteristics of schools which are covered by school size (measured by total number of students), grade, gender and race (from the latter source). Furthermore, students who attempted first time in a school have been considered for calculating the G.C.E O/L pass rate of that school. Samples were made by connecting the above mentioned G.C.E. O/L pass rates of schools data with characteristics of schools data, considering all schools in the Northern and Eastern provinces from 2005 to 2014. Accordingly, the sample size was made as 8,455 school records after cleaning of the data. The G.C.E O/L pass rate of schools is the response variable which is followed by a three level hierarchical data structure. Here, schools are geographically clustered within eight districts while districts are again clustered within the two provinces. In the three-level hierarchical structure, school is represented at level 1, district at level 2 and province at level 3. Table 1 in Appendix I presents variables used in detail of this study for the univariate analysis and advanced analysis.
In the advanced analysis of this study, the Generalized Linear Multilevel Ordinal Model has been developed. This type of multilevel model has the response in the form of an ordered category scale. However, we can see natural ordering and these types of categories can be allocated scores which are treated as if they are measurements of a continuous scale. On some occasions, the distribution of the continuous response variable is skewed. One alternative in such a situation is to discretize the variable into categories using an appropriate method and then assign scores for these categories. In this study, the response variable is continuous and its distribution violates the assumption of a normal distribution. Hence, the response variable G.C.E. O/L pass rate at schools has been categorized using equal width discretization method and then scores were assigned to each category according to Table 1 in Appendix I.
Also, Time period is one of the important factors in this study and it has been divided into two time periods, namely, war period and post-war period. It is important to justify that the individuals are independent between the two time periods. Since the G.C.E O/L pass rate of schools was taken by considering pass rates of students at first attempt in each school, students represented by each school are independent during the two time periods. Accordingly, it is clear that the G.C.E O/L pass rates of schools are independent during war period and post-war period.

2.2 Methodology for Univariate Analysis

The main essence of univariate analysis is to identify whether there is a relationship between each explanatory variable and response variable. It helps to provide a preface to advanced modelling by selecting significant variables among several explanatory variables.
Zhang and Boos [11] introduced a new testing method for the use of correlated categorical data named as the Generalized Cochran Mantel-Haenszel (GCMH) Test. This test provides three different test statistics called T_{EL}, Tp and Tu. According to a simulation study done by Zhang and Boos [11], Tp is considered to be the most appropriate Test Statistics among T_{EL} and Tu for application in multilevel data structures. Zhang and Boos [11] introduced the theory and algorithm for correlated categorical data with repeated measures. Further, the algorithm proposed by Zhang and Boos [11] was modified using the R function by De Silva and Sooriyarachchi [12], in order to apply the test on a two dimensional multilevel dataset without repeated measures. Hence, the most suitable test for univariate analysis for this study is identified as GCMH test which is executed using the developed R function. This function provides output as the value of Tp, the value of GCMH Statistic and the corresponding p values with the degrees of freedom. According to the p value, conclusion can be made as to whether there is a significant association between the response variable and the explanatory variable.

2.3 Methodology for Advanced analysis

Since data used in this study follow a multilevel data structure with a type of ordered categorical response, a Generalized Linear Multilevel Ordinal Model was decided on to develop. Accordingly, this section mentions the theory and methodology behind multilevel model for ordered categorical response described by Rasbash et al. [10]. At the beginning, a single level model with an ordered categorical response variable has been explained as it is generally recommended to begin model building with the simple model. Thereafter, generalization of the two level model by the single level model has been explained.

2.3.1 A Single Level Model with an ordered categorical response variable

According to Rasbash et. al. [10], the following theory for the Single Level Model with an ordered categorical response has been explained.
Suppose $y_{i\ \mathrm{response}}$ variable which is ordered categorical for individual i and it has t categories which is indexed by s, s = 1, 2, t where t is selected as reference category
Here, models with Cumulative Probabilities is considered instead of using response probabilities for each category in order to exploit the ordering.
$E(y_i^{(s)}) = \gamma_i^{(s)} = \sum_{h=1}^{s} \pi_i^{(h)}$ where, **s = 1,2, ……., t-1** ; $\pi_i^{(s)}$ = Probability of i[th] individual having a response variable of s ; $\gamma_i^{(s)}$= The observed Cumulative Proportions (out of the total n_i observations) for the i[th] individual Multinomial logistic model is used and one of the response category is chosen as reference category which is normally taken as the lowest category level in an ordered categorical response model. It is estimated t-1 set of equations for each different response category with the chosen reference category. The proportional odds model with a logit link function is commonly used. Multinomial regression model with explanatory variable X_i can be written in the form of logit link as follows.

$$logit\ (_i^{s}) = ^{(s)} + X_i\beta \quad \text{or} \quad (_i^{s}) = \{\ + \exp -[\alpha^{(s)} + X_i\beta]\}^{-1}$$

It is estimated that a separate intercept and slope exists for each response category. However, a common slope for each response category can be estimated. This indicates that increasing values of the linear component are associated with increasing probabilities as s increases.

2.3.2 A Two-Level Model with an ordered categorical response variable

According to Rasbash et. al. [10] and Fielding et. al. [13], theory for Two-Level Proportional Odds Model with an ordered categorical response and Non-Proportional Odds Model have been explained below.

The two level ordered category response model is a generalization of the single level model. Similar to the single level model, a two level model is considered with Cumulative Probabilities as given below.

$$E(y_{ij}^{(s)}) = \gamma_{ij}^{(s)} = \sum_{h=1}^{s} \pi_{ij}^{(h)};\ s = 1, 2, \dots\dots, t-1$$

Here, y_{ij} is categorical response of individual i in second level j, probability of having category s for i^{th} individual in second level j is denoted by $\pi_i^{(s)}$.

Cumulative proportions have covariance matrix as follows.

$$cov\left(y_{ij}^{(s)}, y_{ij}^{(r)}\right) = \frac{\gamma_{ij}^{(s)}\left(1-\gamma_{ij}^{(r)}\right)}{n_{ij}};\ s \leq r$$

Generalized Multilevel Ordinal Proportional Odds Model for a single random effect and no covariates/factors are mentioned below.

$$logit\ (\gamma_{ij}^{(s)}) = log\left(\frac{\gamma_{ij}^{(s)}}{1-\gamma_{ij}^{(s)}}\right) = \alpha^{(s)} + u_{0j}\ ;s = 1, 2, \dots\dots, t-1$$

It is assumed that the random effect is $u_{0j} \sim (0, \sigma_{u0}^{2})$.

The above model can be extended to fixed or random covariates/factors and several random components as follows.

$$logit\ (\gamma_{ij}^{(s)}) = \alpha^{(s)} + X_{ij}\beta + Z_{ij}u_j\ ;s = 1, 2, \dots\dots, t-1 \quad \text{or}$$

$$\gamma_{ij}^{(s)} = \{1 + \exp -[\alpha^{(s)} + X_{ij}\beta + Z_{ij}u_j]\}^{-1}$$

Where, $X_{ij} = \{1, x_{1ij}, x_{2ij},\}$ and $Z_{ij} = \{1, z_{1ij}, z_{2ij},\}$ while $\beta^T = \{\beta_0, \beta_1,\}$ $U_j^T = \{u_{0j}, u_{1j},\}$

Generally, Z variables are a subset of X variables. β is a vector of fixed effects coefficients associated with the covariates/factors in the data vector X_{ij}. Up $_{is}$ a second level random effect which is assumed to be normally distributed with mean zero and variance σ_u^2. The subjects of U_j^T are random variables at second level j. It is assumed that the individuals of U_j^T are of dependent multivariate normal distribution with mean zero.

Here, the category probabilities can be written in terms of cumulative probabilities as follows.

$$\pi_{ij}^{(h)} = \gamma_{ij}^{(h)} - \gamma_{ij}^{(h-1)}\ ; 1 < h < t\ ,\quad \pi_{ij}^{(1)} = \gamma_{ij}^{(1)}\ ;\ \gamma_{ij}^{(t)} = 1$$

The theory explained above is based on the most basic multilevel model identified for ordinal categorical response. Further, many extensions to this basic model by Fielding et al. [13], Hedeker and Gibbons [14], Raman and Hedeker [15] have been proposed. Accordingly, the extension for the proportional odds model is the generalized multilevel ordinal non-proportional odds model which is generally used when there are substantial reasons to suggest that the effect of explanatory variables do not behave proportionally across response categories.

The following illustrates how the generalized multilevel ordinal non-proportional odds model behaves.

$$logit\ (\gamma_{ij}^{(s)}) = \alpha^{(s)} + \omega^{(s)}t_{ij} + X_{ij}\beta + u_{0j}\ ;s = 1, 2, \dots\dots, t-1$$

Here, ij indicates variables supposed to vary non-proportionally across logits. The terms $\omega^{(s)}$ indicates the estimated coefficients that vary across the logits. Further, Fielding et al. [13] have described these models and their application to educational data.

2.3.3 Variable Selection and Model Comparison

To select significant variables which need to be included into the model, the forward selection procedure was used along with Wald Statistic and Deviance Information Criteria (DIC). The P value corresponding to Wald Statistics was used to find out the most significant variable that needed to be included into the model while DIC was used to evaluate the model fit at each stage of the forward selection procedure.

2.3.4 Deviance Information Criteria (DIC)

The deviance statistic is known as a measure of how well a developed model fits the data [16]. Here, deviance is calculated by running MCMC to derive a diagnostic which is known as Deviance Information Criteria. This DIC diagnostic is a generalization of the Akaike's Information Criteria (AIC) [17]. Consider P_D is an effective number of parameters, $(\bar{\theta})$ is deviance at the expected value of the unknown parameters θ and $\bar{D}$ is the average deviance from the complete set of iterations. Then, DIC = $\bar{D} + P$ where $P_D = \bar{D} - D(\bar{\theta})$.

Since the DIC diagnostic has two terms which explain fits of a model and its complexity, DIC helps to compare models. Generally, the smaller DIC value represented model is considered as a better model.

2.3.5 Parameter Estimation Procedure

Since maximum likelihood estimation is not applicable to the discrete multilevel response model, Quasi-likelihood methods are executed in MLwiN 2.10. These Quasi-likelihood methods use a linearization method which is based on Taylor Series Expansion. The Taylor Series Expansion transforms a discrete response model into a continuous response model. Also, there are two types of approximation methods available in MLwiN 2.10 to transform a model into a linear model. These types of methods are known as Marginal Quasi Likelihood (MQL) and Predictive (or Penalized) Quasi-likelihood (PQL) and they are used with 1st order or 2nd order terms in the Taylor series expansion. After linearization, estimation can be executed using IGLS or Reweighted IGLS (RIGLS) in MLwiN software.

The IGLS and RIGSL methods can be identified as likelihood based frequentist methods and they are founded on iterative procedures. These iterations happen between two deterministic steps till two consecutive estimates for each parameter are provided as sufficiently close together. However, estimates produced through these methods tend to be biased.

Hence, instead of the above mentioned two methods for parameter estimation, recently the MCMC method based on simulation procedures has been implemented and it has several advantages compared to IGLS and RIGLS methods. Instead of providing point estimates for unknown parameters, these methods are used to run many iterations and to create estimate for each parameter at each iteration. Also, these estimates are not independent since at each iteration, the estimates at the last iteration are used to create new estimates. MCMC estimation creates sample values from the posterior distribution of the unknown parameters and this helps to provide accurate estimates for the unknown parameters of the model. Also, it allows to calculate point and interval estimates for each level considered in the multilevel model [1]. However, MCMC estimation is computationally intensive and it takes a far longer time to run than likelihood based IGLS and RIGLS methods using MLwiN.

In this study, the IGLS method with either PQL 1st order or PQL 2nd order method is firstly used to find starting estimates for model parameters. Thereafter, the MCMC method which is known as a good estimation method for parameter estimation is executed to find out better estimates for each parameter in the model.

2.4 Residual Analysis and Model Adequacy

After fitting a model to a set of data, it is important to check whether model fitting is adequate or not. The general way of finding the model adequacy is based on residuals analysis which is aimed to find out whether the underlining distributional assumptions of error terms are validated by the fitted model. If it is true, then the fitted model can be used to make valid inferences. It could be recognized that the researchers have not addressed up to now, the area of the residual analysis and diagnostic tests for multilevel modelling thoroughly. Hence, the available methods have been used to validate the fitted model.

Even though the multilevel ordered categorical response model has been developed in this study, diagnostic methods used for model adequacy are the same as other types of multilevel models. Accordingly, the theory applied for evaluation of a model adequacy under a multilevel model with continuous response by Rasbash et al. [10] has been used in this study.

Generally, it is found that the residual analysis using higher level aggregation rather than observing individual data points is more useful. Reasons for this choice is that most often researchers prefer to take higher level aggregation as they normally tend to be started at this level and when differences at the higher level exists, higher level indication is clearer than observing a few individual points at a lower level [10].

2.4.1 Methods used in Checking Model Adequacy

Generally, once estimated residuals are produced at any level, the residuals at any level follow a normal distribution and can be checked using a normal probability plot. Further, the Anderson Darling test is used to check the normality using Minitab 16.

In multilevel modelling, caterpillar plot can be used to illustrate and compare random effects parameters. It can be drawn by plotting residuals at any level in ascending order with their 95% confidence interval using MLwiN 2.10. This plot is displayed as a caterpillar as its name implies. Here, individuals of the considered level are generally shown at the lower and upper sides of the plot and the confidence interval of their residuals do not overlap zero. These residuals shows that the deviation of individuals at any level from the overall mean predicted by the fixed parameter in the model. That means these are individuals that indicate significant changes from overall mean at 5 % level of significance [10].

3. Results

3.1 Results of Univariate Analysis based on the Generalized Cochran Mantel-Haenszel test

Mainly, univariate analysis was conducted to identify whether there is an association between each explanatory variable and response variable the G.C.E. O/L pass rate. Also, significant variables obtained in the univariate analysis would be helpful to make a comparison with significant variables obtained in the final model under the advanced analysis. In this study, data used follow a three level hierarchical structure and initially there are five explanatory variables which are types of categorical variables except the school size variable. As mentioned in the methodology for univariate analysis, the GCMH test used under the univariate analysis requires categorical data in a two level hierarchical data structure. Therefore, district level 2 and province level 3 were combined to make a single variable prior to running the R function. Also, initially explanatory variable school size is a type of continuous variable and it was categorized meaningfully by referring to the Sri Lanka Education Information 2013 report [18]. The following table illustrates output of the GCMH test which was executed using the developed R function by De Silva & Sooriyarachchi [12].

Table 2: Generalized Cochran Mantel-Haenszel test results on Explanatory Variables with G.C.E. O/L Pass Rate

Explanatory Variable	Tp	Degrees of Freedom	P-Value
Time Period	128.294	3	1.261E-27*
Grade	1190.920	6	4.418E-254*
Race	246.450	6	2.352E-50*
Gender	359.037	6	1.771E-74*
School Size	1026.244	27	6.447E-199*

* Significant at 0.01% Level

H_o: There is no association between G.C.E. O/L Pass Rate and Explanatory Variable Vs.
H_1: There is an association between G.C.E. O/L Pass Rate and Explanatory Variable
According to results of GCMH test presented in Table 2, since P value for each explanatory variable is less than 0.0001, H_o is rejected at 0.01% level of significance. It would imply that all variables are significantly associated with G.C.E. O/L Pass Rate. Moreover, the highest significance has been indicated from variable Grade.

3.2 Model Building Procedure for Generalized Linear Multilevel Ordinal Model for G.C.E O/L Pass Rate

3.2.1 Generalized Linear Multilevel Ordinal Model with three levels

This section describes advanced modelling of response G.C.E. O/L pass rate with explanatory variables using MLwiN 2.10 software. All explanatory variables have been taken into account for model building based on results of the GCMH test under univariate analysis.

The useful strategy in multilevel model similar to other statistical models, is to start by fitting a simple model and gradually increasing the complexity. Initially, a generalized linear multilevel model for ordered categorical response with three levels (i.e., level 1 is School, level 2 is District and level 3 is Province) was developed under two approaches named as the proportional odds model and the non-proportional odds model. The forward selection method was adopted for the model building procedure and it was started from the basic variance component model. Accordingly, each explanatory variable was added to the basic model one at a time in order to fit the final main effects model. It is important to mention, that in the forward selection procedure, once a variable is selected to the model, the selected variable should not be ignored throughout the process even when that variable is insignificant at proceeding stages. When fitting the model at each stage, either the 1st order or 2nd order PQL method was used based on convergence. Thereafter the MCMC method was used by running up to 5,000 iterations to find out better estimate for each parameter in the model.

A better model at each stage was selected considering a model with minimum DIC value together with identifying significant categories of a variable included into the model using p value of the Wald statistics. Accordingly, both proportional and non-proportional odds models were built separately in order to find the most appropriate final main effects odds model. Here, DIC value obtained for the main effects Propotional odds model was 19,614.00 while DIC value for the main effects Non-propotional odds model was 19,363.35. Since the minimum DIC value is represented by the final main effects on the non-proportional odds model, it was selected as the most appropriate final main effects odds model.

After fitting a multilevel model, it is essential to justify whether the fitted multilevel model is adequate or not. To test the availability of multilevel for the selected model, province and district level variations were checked using 95% confidence interval. If zero is included within the 95% confidence interval, H_0 is not rejected. If H_0 is not rejected at both levels, it would imply that multilevel is not applicable since it is the same as a single level data analysis.

Table 3: Estimate of province and district levels variances for fitted Main Effects Non-Proportional Odds Model

Parameter	Estimate (95% Confidence Interval)
Unexplained Province Level Variance	0.055 (0.000,0.471)
Unexplained District Level Variance	0.332 (0.110,0.916)

The following hypothesis is used for testing the Province level (or District Level) variations.
H_0: Unexplained Province Level (or District Level) variance is zero Vs.
H_1: Unexplained Province Level (or District Level) variance is not zero

According to results of Table 3, it is revealed that the estimated value for variance of the third level province for the selected model is not significant at 5% level of significance while the estimated value for variance of the second level District for the selected model is significant at 5% level of significance. Hence, it was decided to remove the third level

Province from the model and to develop a Generalized Linear Multilevel Model with only two levels (i.e., level 1 is School, level 2 is District).

3.2.2 Generalized Linear Multilevel Ordinal Model with two levels

Similar to the above mentioned model building procedure, a Generalized Linear Multilevel Ordinal Model with two levels (i.e., level 1 is School, level 2 is District) was developed under two approaches, namely, the proportional odds model and the non-proportional odds model using the forward selection method. However, under the non-proportional odds model approach, when School size was included as a separate coefficient into the model, MLwiN 2.10 crashed. Hence, the only school size variable had to be added to the model as a common coefficient at each step of the forward selection procedure. Accordingly, a proportional odds model and a partial non-proportional odds model were developed separately. The DIC value obtained for the propotional odds model was 19,614.29 while DIC value for the partial non-propotional odds model was 19,397.27. Since the minimum DIC value was shown by the final main effects partial non-proportional odds model, it was identified as the better main effects model than the propotional odds model.

In order to justify whether the selected final main effects partial non-proportional odds model is adequate for multilevel analysis, the district level variation was checked using 95% confidence interval as previously described in section 3.2.1. According to the 2.5th percentile value of 0.098 and 97.5th percentile value of 0.894 for the district level variance, it concludes that the above interval does not include zero and therefore the null hypothesis (H_0) is rejected at 5 % level of significance. It would imply that the district level variation is present in the selected final main effects partial non-proportional odds model.

3.3 Results for Residual Analysisand Model Adequacy

It is important to check adequacy of the fitted final model in order to make accurate and valid inferences from the model. Hence, as mentioned inspection 2.4.1, two graphical techniques named the Caterpillar plot and Normal probability plot were used to check the adequacy of the fitted final main effects partial non-proportional odds model.

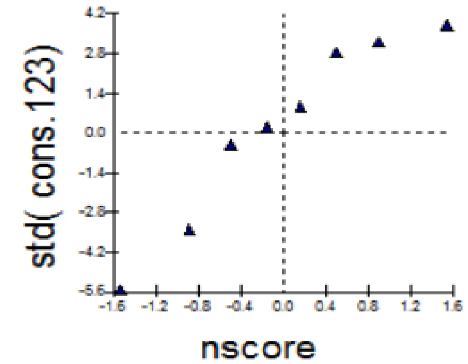


Fig 1: Normal Probability Plot for final Main Effects Partial Non-Proportional Odds Model

Table 4: Anderson Darling Test for Final Main effects Partial Non-Proportional Odds Model

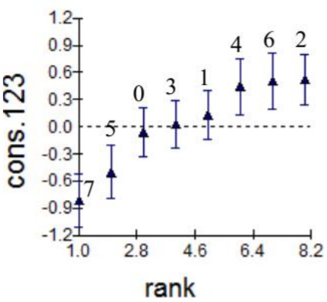
AD Statistic	0.092
P-Value	0.995

According to Fig 1: Normal Probability Plot, most of the points lie on the line angle at 45^0. Hence, it can be reasonably concluded that the residuals satisfy the normality assumption by eye inspection. However, the Anderson Darling Test was carried out to obtain a precise conclusion for adequacy of this fitted Model. The Hypothesis for the Anderson Darling Test is as follows.

$H_{0:}$ The residuals follow normal distribution Vs.

$H_{1:}$ The residuals do not follow normal distribution

According to test results of table 4, since p-value for Anderson Darling Test statistic is 0.995 and it is greater than 0.05, it is not feasible to reject H_0 at 5 % level of significance. That means the residuals are distributed normally.



Abbreviation used for Districts;

0 = Ampara
1 = Batticaloa
2 = Trincomalee
3 = Jaffna,
4 = Killinochchi
5 = Mannar
6 = Mullativu
7 = Vavuniya

Fig 2: Caterpillar Plot for final Main Effects Partial Non-Proportional Odds Model

According to Fig 2: Caterpillar Plot, it clearly indicates that the 95% confidence intervals for five districts Vavuniya, Mannar, Kilinochchi, Mullativu and Trincomalee have not included zero while 95% confidence intervals for three districts Ampara, Batticaloa and Jaffna have included zero. Accordingly, it implies that the performance at G.C.E. O/L in Vavuniya and Mannar districts is worse than the overall mean performance at G.C.E. O/L while performance at G.C.E. O/L in Kilinochchi, Mullativu and Trincomalee districts is better than the overall mean performance at G.C.E. O/L.

3.4 Interaction Model

After selecting the most appropriate main effects odds model, it was decided to proceed with model building by adding interaction terms into the selected main effects model because it is known that adding interactions between main effects to a model would result in a better model. Here, the forward selection method was adopted to select the most significant interactions for the main effects model while interaction terms were included into the model as a separate coefficient or common coefficient according to appropriateness. Accordingly, DIC value of the final Interaction Model was obtained as 19,315.96 which is the lowest DIC value when compared to DIC values of the other fitted models. However, when considering the final interaction model, it has twenty five significant categories of variables and eleven significant interaction terms.

Generally, **s**implicity is an important feature of a model since it is easy to understand and make inferences from the model. This is known as the Parsimony Principal. Also, a parsimonious model provides better predictions than one which is very complicated. Accordingly, when considering the final interaction model, it is a complicated and complex model to do the interpretations. On the other hand, DIC value of the selected final main effects partial nonproportional odds Model was 19,397.27 and the difference between DIC values of the final main effects partial non-proportional odds model and final interaction model is 81.31. Also, this difference can be considered as smaller when compared to the difference between DIC values of the proportional odds model and final main effects partial non-proportional odds model with two levels, which is 217.02. Based on these reasons, it was decided to select the final main effects partial non-proportional odds model with two levels (i.e., level 1 is School, level 2 is District) as the most appropriate model for response, G.C.E O/L pass rate.

Final Main Effects Partial Non-Proportional Odds Model

$$resp_{ijk} \sim \text{Ordered Multinomial}\,(cons_{jk}, \pi_{ijk})$$

$$\gamma_{1jk} = \pi_{1jk}, \gamma_{2jk} = \pi_{1jk} + \pi_{2jk}, \gamma_{3jk} = \pi_{1jk} + \pi_{2jk} + \pi_{3jk}, \gamma_{4jk} = 1$$

$$logit\,(\gamma_{1jk}) = -3.313\,(0.258)\,cons.(<= 0 - 25.00)_{ijk} - 2.648(0.277)\ 1AB\ school.(<= 0 - 25.00)_{ijk} - 1.115(0.092)1C\ school.(<= 0 - 25.00)_{ijk} - 2.068(0.466)\,Girls'school.(= 0 - 25.00)_{ijk} - 0.561(0.067)Post - warperiod.(<= 0 - 25.00)_{ijk} + h_{jk}$$

$$logit\,(\gamma_{2jk}) = -1.202(0.254)\,cons.(<= 25.01 - 50.00)_{ijk} - 1.454(0.094)1AB\ school.(< 25.01 - 50.00)_{ijk} - 0.411(0.059)1C\ school.(<= 25.01 - 50.00)_{ijk} - 1.160(0.142)Girls'\ school.(<= 25.01 - 50.00)_{ijk} - 0.558(0.049)Post - warperiod.(\leq 25.01 - 50.00)_{ijk} - 0.310(0.094)\ Muslim\ school.(<= 25.01 - 50.00)_{ijk} + h_{jk}$$

$$logit\,(\gamma_{3jk}) = 0.623(0.258)\,cons.(<= 50.01 - 75.00)_{ijk} - 0.558(0.104)\ 1AB\ school.(< 50.01 - 75.00)_{ijk} - 1.332(0.113)Girls'\ school.(<= 50.01 - 75.00)_{ijk} - 0.494(0.064)\ Post - warperiod.(\leq 50.01 - 75.00)_{ijk} - 0.865(0.113)Muslim\ school.(<= 50.01 - 75.00)_{ijk} + h_{jk}$$

$$h_{jk} = 1.364(0.485)\ 1 - 50.123_{jk} + 2.182(0.282)\ 51 - 100.123_{jk} + 2.282(0.267)\,101 - 200.123_{jk} + 2.161(0.267)\,201 - 500.123_{jk} + 1.832(0.273)\,501 - 1000.123_{jk} + 1.530(0.287)1001 - 1500.123_{jk} + 0.899(0.303)1501 - 2000.123_{jk} + v_{4k}cons.123$$

$$v_{4k} \sim N\,(0, \Omega_v) : \Omega_v = [0.308(0.220)]$$

$$cov\,(y_{sjk}, y_{rjk}) = \gamma_{sjk}\,(1 - \gamma_{rjk})/\,cons_{jk}\ ;\quad s \leq r$$

γ_{1jk} = Probability that school j clustered within district k, having at most 25.00 % G.C.E. O/L pass rate
γ_{2jk} = Probability that school j clustered within district k, having at most 50.00 % G.C.E. O/L pass rate
γ_{3jk} = Probability that school j clustered within district k, having at most 75.00 % G.C.E. O/L pass rate

The above illustrates the final main effects partial non-proportional odds model for G.C.E O/L pass rates with only significant categories of variables. It is followed by three response variable equations, one for each cumulative category. The model consists of unique coefficients for every response category as well as a common coefficient. Here, significant categories of school size variable are represented as common coefficient for each logits while significant categories for other explanatory variables named as Grade, Gender, Time period and Race indicate unique coefficients for each logit. The first explanatory variable in each category of response variable is a constant. The other explanatory variable, $cons.$ 123 is also constant which is equal to one, whose sole distribution to the model through its random coefficients is to add the same random error term to each of the three categories equations. Here, term h_{jk} indicates terms common to the set of equations for each G.C.E. O/L Pass Rate response category. Also, $cons.$ 123 is used as coefficient of the common variable to specify as a common between district level variability.

Table 5: Final Main Effects Odds Model with separate coefficients of varaibles

Variable	Category	≤ (0 - 25.00)			≤ (25.01 - 50.00)			≤ (50.01 - 75.00)		
		$\hat{\beta}$	SE($\hat{\beta}$)	P-Value	$\hat{\beta}$	SE($\hat{\beta}$)	P-Value	$\hat{\beta}$	SE($\hat{\beta}$)	P-Value
Grade	1 AB School	-2.648	0.277	1.183E-21*	-1.454	0.094	5.697E-54*	-0.558	0.104	8.079E-08*
	1 C School	-1.115	0.092	8.319E-34*	-0.411	0.059	3.258E-12*	0.133	0.085	0.118
Gender	Boys' School	0.385	0.229	0.093	-0.139	0.158	0.379	-0.273	0.166	0.100
	Girls' School	-2.068	0.466	9.089E-06*	-1.160	0.142	3.109E-16*	-1.322	0.113	1.288E-31*
Time Period	Post-war Period	-0.561	0.067	5.612E-17*	-0.558	0.049	4.812E-30*	-0.494	0.064	1.175E-14*
Race	Tamil School	0.08	0.1	0.424	0.075	0.082	0.360	-0.156	0.106	0.141
	Muslim School	0.077	0.129	0.551	-0.310	0.094	9.742E-04*	-0.865	0.113	1.935E-14*

* Significant at 5% Level

Table 6: Final Main Effects Odds Model with common coefficients of varaible

Variable	Category	$\hat{\beta}$	SE($\hat{\beta}$)	P-Value
School Size	1-50	1.364	0.485	4.92E-03*
	51-100	2.182	0.282	1.013E-14*
	101-200	2.282	0.267	1.265E-17*
	201-500	2.161	0.267	5.791E-16*
	501-1000	1.832	0.273	1.94E-11*
	1001-1500	1.530	0.287	9.77E-08*
	1501-2000	0.899	0.303	3.01E-03*
	2001-2500	0.259	0.332	0.435
	2501-3000	0.244	0.394	0.535

* Significant at 5% Level, DIC Value = 19,397.27

Table 5 and Table 6 illustrate p-values of Wald statistics for each category of variables in the selected final main effects odds model. According to Table 5, all categories of the variable Grade in all three logits except 1C school category in (≤ 50.01 - 75.00) logit, are significant at 5% level. Also, the variable Time Period for all three logits are significant at 5% level. In the variable Gender, Boys' school category for all three logits are not significant at 5% level while Girls' school category for all three logits are significant at 5% level. In the variable Race, Tamil school category for all three logits and Muslim school category for only (≤ 0-25.00) logit are not significant at 5% level while Muslim school category for the other two logits are significant at 5% level. Moreover, table 6 shows that all categories other than (2001-2500) and (2501-3000) categories of the variable school size are significant at 5% level.

3.5 Results of final main effects partial non-propotional odds model

It is essential to interpret the parameter estimates of a model, once a model is fitted and its adequacy is verified. Hence, the following tables 7 and 8 illustrate the odds ratios of significant categories of each variable of the final main effects odds model for the response, G.C.E. O/L pass rate.

Table 7: Odds ratios of significant categories of each variable as separate coefficients in the Final Main Effects Odds Model

Variable	Category	Odds ratio		
		≤ (0 - 25.00)	≤ (25.01 - 50.00)	≤ (50.01 - 75.00)
Grade	1 AB School [Type 2 School [b]]	0.071	0.234	0.572
	1 C School [Type 2 School [b]]	0.328	0.663	-
Gender	Girls' School [Mixed School [b]]	0.126	0.313	0.264
Time Period	Post-war Period [War Period[b]]	0.571	0.572	0.610
Race	Muslim School [Sinhala School [b]]	-	0.733	0.421

[b] = Base Category

Table 8: Odds ratios of significant categories of each variable as common coefficients in the Final Main Effects Odds Model

Variable	Category	Odds ratio
School Size	1-50	3.911
	51-100	8.864
	101-200	9.796
	201-500	8.680
	501-1000	6.246
	1001-1500	4.618
	1501-2000	2.457

Base Category = [3001-3500]

In Table 7, odds ratios are calculated and presented for each significant category of variables Grade, Gender, Time Period and Race since these variables have been fitted as separate coefficients in the final main effects odds model. Also, table 8 shows common odds ratio for each significant category of variable school size since school size has been fitted as common coefficient in the final main effects odds model. Following sections will discuss the impact of all significant terms of the final main effects odds model for response of G.C.E. O/L pass rate.

According to the results in table 7, odds ratio of having a pass rate ≤ 25.00 % as opposed to a pass rate > 25.00 % for 1 AB Schools is 0.07 times compared to Type 2 schools while 1C Schools are 0.33 times these odds. Therefore, a significant difference can be noticed between the odds ratios of 1 AB schools and 1C schools with respect to the pass rate category ≤ (0-25.00)

%. The odds ratio of having a pass rate ≤ 50.00 % as opposed to a pass rate > 50.00% for 1 AB Schools is 0.23 times compared to Type 2 Schools while it is 0.66 times for 1 C schools. Accordingly, 1 AB schools have the lowest odds of having a G.C.E. O/L pass rate ≤ 50.00 % as opposed to a G.C.E. O/L pass rate > 50.00 % when compared to 1C schools and Type 2 schools.

Also, the odds ratio of School having a pass rate ≤ 25.00 % as opposed to pass rate > 25.00 % for Girls' schools is approximately 0.13 times compared to Mixed schools while odds ratio of having a pass rate ≤ 50.00 % as opposed to a pass rate > 50.00 % for Girls' Schools is 0.31 times compared to Mixed Schools. The odds ratio of having a pass rate ≤ 75.00 % as opposed to pass rate > 75.00 % for Girls' schools is 0.26 times compared to Mixed schools. Accordingly, odds ratio of having a pass rate ≤ 50.00 % as opposed to a pass rate > 50.00 % for Girls' schools is slightly equal to odds of having a pass rate ≤ 75.00 % as opposed to a pass rate > 75.00 % for Girls' schools.

The odds ratio of school having a pass rate ≤ 25.00 % as opposed to a pass rate > 25.00 % during the post-war period is approximately half times higher compared to the war period. The odds ratio of a school having a pass rate ≤ 50.00 % as opposed to a pass rate > 50.00 % during the post-war period is approximately half times higher compared to the war period. Also, the odds of schools having a pass rate ≤ 75.00 % as opposed to a pass rate > 75.00 % during the post-war period is approximately 0.6 times compared to the war period. Accordingly, it indicates that odds of schools having a less than or equal G.C.E O/L pass rate category as opposed to a greater G.C.E O/L pass rate category during the post-war period is significantly lower compared to the war period.

Moreover, the odds ratio of schools having a pass rate ≤ 50.00 % opposed to a pass rate > 50.00 % for Muslim schools is 0.73 times compared to Sinhala schools. Also, the odds of schools having a pass rate ≤ 75.00 % as opposed to a pass rate > 75.00 % for Muslim schools is approximately 0.42 times compared to Sinhala schools. Accordingly, it would imply that odds of schools having a less than or equal G.C.E O/L pass rate category as opposed to a greater G.C.E O/L pass rate category for Muslim schools is lower compared to Sinhala schools.

According to table 8, the odds ratio of having less than or equal pass rate category as opposed to a greater pass rate category (when considering all three logits) for schools with size (1-50) range is nearly 4.0 times compared to schools with size (3001-3500) range. Similarly, school sizes within ranges (51-100), (101-200), (201-500), (501-1000), (1001-1500) and (1501–2000) have shown approximately 8.9, 9.8, 8.7, 6.2, 4.6 and 2.5 times these odds respectively. Accordingly, the noticeable difference of corresponding odds ratios of schools sizes with ranges (101-200) and (1501-2000) can be observed. Moreover, when school size is increased from (101-200) range to (1501-2000) range, their corresponding odds ratios have decreased.

4. Discussion

Multilevel modelling is a widely used area for hierarchical data structures. The analysis of ordered categorical response followed by multilevel data structure is an interesting approach as well as a new application for the study area. Also, multilevel modelling for Educational data has been hardly applied in Sri Lanka when compared to other countries. Hence, this study would be a novel statistical application based on Education and conflict in Sri Lanka.

MLwiN 2.10 software is a powerful statistical software for developing multilevel models as it facilitates easy method for inclusion of variables and interaction terms into a model at any levels of the hierarchy. The importance of considering a multilevel structure for G.C.E. O/L pass rate is that it eliminates misleading results such as obtaining biased estimates and large standard errors. Also, the multilevel modelling technique explains how the response variable behaves with the multiple levels simultaneously.

An acceptable method for the advanced analysis has been followed in this study. MLwiN 2.10 has been used for fitting the Generalized Linear Multilevel Model for ordered categorical response. The MCMC estimation method was used up to 5000 iterations after running the 1st order or 2nd order PQL method for parameter estimation in the models. Initially, a multilevel model with three levels (i.e., level 1 is School, level 2 is District and level 3 is Province) could be developed under two approaches, namely, the proportional odds model and the nonproportional odds model. Further, it is proved that the province level is not adequate for fitting the model with a three level hierarchical structure. Accordingly, a Generalized linear multilevel model with two levels (i.e., level 1 is School, level 2 is District) could be developed similar to the model with three levels. At many instances, MLwiN 2.10 crashed while running the nonproportional odds model with adding the school size variable as a separate coefficient into that model using MCMC estimation. Accordingly, the partial non-proportional odds model with two levels could be developed instead of building the non-proportional odds model. Although the interaction model was built after fitting the final main effects odds model and its DIC value was minimum, the interaction model was a complex model to do the interpretations. Accordingly, there is a tradeoff between accuracy and simplicity when selecting the most appropriate final model. Ultimately, the partial non-proportional odds model with two levels was selected as the most appropriate final main effects odds model for the response, G.C.E. O/L pass rate based on its simplicity and accuracy. A sound multilevel goodness of fit test has not been recommended for the multilevel models. Hence, adequacy of the final main effects odds model has been checked using the Caterpillar plot and the Normal probability plot.

This paper is mainly focused on how conflict affects the G.C.E. O/L pass rate of schools in the Northern and Eastern provinces of Sri Lanka. Through this study, it was possible to obtain good results based on the developed final main effects odds model. A careful interpretation for odds ratios in the final main effects partial non-proportional odds model is important since these are somewhat difficult to interpret. As explained under the results section, the variable time period was significant at 5% level of significance while not all categories of other variables Grade, Gender, Race and school size were significant at 5% level of significance in the final main effects odds model. Also, it is proved that the odds of schools having a less than or equal pass rate as opposed to a greater pass rate during the post-war period is significantly lower compared to the war period. Accordingly, it is concluded that there is a significant improvement of G.C.E.

O/L pass rate of schools during the post-war period compared to the war period in the Northern and Eastern provinces of Sri Lanka. Moreover, 1 AB schools indicate the highest G.C.E. O/L performance when compared to the 1C and Type 2 schools. Also, the G.C.E. O/L performance of Girls' schools is higher than that of mixed schools while the G.C.E. O/L performance of Muslim schools is higher than that of Sinhala schools in Northern and Eastern provinces of Sri Lanka.

Moreover, it is proved that the performances at the G.C.E. O/L Examination in Vavuniya and Mannar Districts are worse than the overall performance at the G.C.E. O/L while performances at the G.C.E. O/L in Kilinochchi, Mullativu and Trincomalee Districts are better than the overall performance at the G.C.E. O/L. Also, Ampara, Batticaloa and Jaffna performances at the G.C.E.
O/L are similar to the level of overall performance at the G.C.E. O/L.

4.1 Further Study

The data used in this study was limited to the Northern and Eastern provinces based on the main objective of this study. It is recommended to apply the technique used here to a sample of schools drawn from all provinces in Sri Lanka. It would be useful to find out how the progress of the G.C.E. O/L pass rates are in the most affected provinces compared to the other provinces in Sri Lanka. In a statistical context, there may be some multi-collinearity variables in the data set. However, it was not considered in this study. Hence, an appropriate test can be applied to find multi-collinearity between variables.

5. Acknowledgement

The authors wish to profoundly thank the Department of Examinations and the Data Management Branch of the Ministry of Education for providing the much required data for the study.

6. References

[1]. Brown WJ, Rasbash J, Charlton C. MCMC estimation in MLwiN Version 2.10; 2009.

[2]. Kuruppumullage P and Sooriyarachchi MR. Log linear models for ordinal multidimensional categorical data. Journal of the National Science Foundation of Sri Lanka. 2007; 35 (1), 29– 40.

[3]. Udaya Kantha KKS and Perera KED. Accelerating Inclusive Growth Trends in Sri Lanka: Challenges and Achievements. Social Sciences and Humanities Review; 2013.

[4]. Valente C. Education and civil conflict in Nepal. World Bank Economic Review. 2014; 28(2), 354–383.

[5]. United Nations Children Fund (UNICEF). Colombo, Sri Lanka: Out-of-School Children in Sri Lanka: Country Study; 2013.

[6]. Goldstein H. Multilevel Statistical Models. *London: Institute of Education*. Multilevel Models. Project, April 1999; 1999.

[7]. Aitkin M, Anderson D, Hinde J. Statistical modelling of data on teaching styles (with discussion). Journal of the Royal Statistical Society. 1981; 144 (A), 148 – 61.

[8]. Aitkin M and Longford N. Statistical modelling in school effectiveness studies (with discussion). Journal of the Royal Statistical Society. 1986; 149 (A), 1– 43.

[9]. Goldstein H, Rasbash J, Yang M, Woodhouse G, Pan H, Nuttall D, Thomas S. A multilevel analysis of school examination results. Oxford review of Education. 1993; 19, 425-33.
[10]. Rasbash J, Steele F, Browne JW, Goldstein H. *A User's Guide to MLwiN*; 2012.
[11]. Zhang J and Boos DD. Generalized Cochran Mantel-Haenszel Test Statistics for correlated categorical data. Communications in Statistics - Theory and Methods. 1997;26 (8), 1813– 1837.
[12]. De Silva DBUS and Sooriyarachchi MR. Generalized Cochran Mantel Haenszel test for multilevel correlated categorical data: an algorithm and R function. Journal of the National Science Foundation of Sri Lanka. 2012; 40(2).
[13]. Fielding A, Yang M, Goldstein H. Multilevel ordinal models for examination grades. Statistical Modelling. 2003; 3(2), 127–153.
[14]. Hedeker D and Gibbons RD. A random-effects ordinal regression model for multilevel analysis. Biometrics. 1994; 50 (4), 933 – 944.
[15]. Raman R and Hedeker DA. Mixed-effects regression model for three level ordinal response data. Statist. Med. 2005; 24, 3331–3345, 2005.
[16]. McCullagh P and Nelder JA. *Generalized linear models* (Vol. 37). CRC press; 1989.
[17]. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2002; 64 (4), 583-639.
[18]. Data Management Branch, Ministry of Education, Sri Lanka: Sri Lanka Education Information; 2013.

Appendix I

Table 1: Description of the data used for Univariate analysis and Advanced analysis

Variable	Description	Category	Code
Cluster Levels			
Province	Level 3 – Province	Eastern	0
		Northern	1
District	Level 2 – District	Ampara	0
		Batticaloa	1
		Trincomalee	2
		Jaffna	3
		Kilinochchi	4
		Mannar	5
		Mullativu	6
		Vavuniya	7
Response Variable			
G.C.E. O/L Pass Rate	G.C.E. O/L Pass Rate of School is calculated as (number of qualified student / number of sat student) *100	0-25.00	1
		25.01-50.00	2
		50.01-75.00	3
		75.01-100.00	4
Explanatory Variables at School Level			
Race		Sinhala School	1
		Tamil School	2
		Muslim School	3
Gender	School categorized based on Gender	Boys' School	1
		Girls' School	2
		Mixed School	3
Grade	1 AB: Schools having Grade 1 to 13 with the addition of a Science stream in Grades 12-13	1 AB School	1
	Type 1C: Schools having Grade 1 to Grade 13, with Arts and Commerce stream in Grades 12-13	1 C School	2
	Type 2: Schools having class up to Grade 11	Type 2 School	3
Time Period	Year range as 2005-2014	War Period (2005-2009)	0
		Post-war Period (2010-2014)	1
School Size	Total number of students in a School	1-50	1
		51-100	2
		101-200	3
		201-500	4
		501-1000	5
		1001-1500	6
		1501-2000	7
		2001-2500	8
		2501-3000	9
		3001-3500	10