

The Effect of Sample Size on Random Component in Multilevel Modeling

Asadullah*, Md. Maidul Husain

Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University,

Gopalganj-8100, Bangladesh

asadullahstat@gmail.com, mhusain@isrt.ac.bd

*Correspondence E-mail: asadullahstat@gmail.com

ABSTRACT. In cluster-correlated data arise when there exists any condition for that individual are grouped among themselves. Data of this kind arise frequently in social science, behavioral, and medical sciences since individuals can be grouped in so many different ways. Multilevel modeling (MLM) is an approach that can be used to handle cluster or grouped data. Analyzing of correlated data is different from the usual way for independent data since we have to consider the correlation structure among individuals within cluster. In random effects models' correlation structure can be estimated by considering the models parameters are allowed to vary across the cluster. Random effect models have two components, within cluster components, cluster-specific response is described by a regression model with a population-level intercept and slope, other is between-cluster component: variation in cluster-intercepts and slopes is captured. In a multilevel model, cluster level variance component is more affected by no. of cluster as well as cluster size. So, this is important to aware the researcher about no. of cluster and cluster size in estimating the random components of random effect models for correlated continuous and discrete outcome respectively in MLM since it produces bias estimate for few no. of cluster and cluster size. The parameters of random effects models can be estimated by Maximum Likelihood and Restricted Maximum Likelihood (REML) estimation for correlated continuous outcome, on the contrast besides REML, Penalized Quasi Likelihood (PQL) and Adaptive Gaussian Quadrature (AGQ) estimation techniques are applied for correlated discrete outcome. In this thesis, using the simulation procedure we would be compared among these estimation techniques by exploring the influence of no. of cluster and cluster size on estimated random parameters from random effect models of two-level and three levels for continuous and discrete outcome respectively. Relative bias, mean square error and coverage probability would be used for comparison purpose among the estimation techniques.

Received: 8 Jan 2022.

Key words and phrases. cluster-correlated data; MLM; random effect models; cluster level; correlation.

In a summary, we find that if we increase the no. of cluster and cluster size at each level biasness and mean square error of estimated random parameter would be decreased for all estimation techniques.

1 Introduction

Clustered data, where the observations are group or cluster. Each cluster contains multiple observations, giving the data a “nested” or “hierarchical” structure, with individual observations nested within the cluster. The key feature of clustered data is that observations within a cluster are “more alike” than observations from different clusters. In clustered data, the dependent variable is measured once for each subject, but the subjects themselves are somehow grouped. The fact that observations within a cluster are more alike than observations from different clusters induces a correlation between observations within the same cluster. This is referred to as intra cluster correlation.

For example, consider the school example; in here 10 schools and each school have 20 students. If we collected data from students from 10 schools then it is cluster data. In this example 10 schools is consider as cluster and the number of students is considered as individuals. Here each individual within correlated in each cluster but uncorrelated between cluster. In multilevel models, however, there is a sample size for each level, defined as the total number of units observed for this level. For example, in a three-level study of pupils nested in classrooms nested in schools, there might be observations on 60 schools, a total of 150 classrooms, and a total of 3,300 pupils.

When data are sampled in a multi-stage manner or if observations are clustered, modeling data by ignoring the clustering will often result in standard error estimates that are underestimated if the outcome variable demonstrates dependence based on the clustering (i.e., the intraclass correlation is greater than zero). When clustering is ignored, the residuals will not be identically and independently distributed, violating an assumption of single-level models such as ordinary least-squares regression. This dependence will ultimately result in an inflated type-I error rate for significance tests of regression coefficients. However, in the statistical literature, methods have been developed for addressing data that come from a hierarchical structure and can account for the

dependence among observations. Clustered data are characterized as data that can be classified into a number of distinct groups or “clusters” within a particular study. However, there are many other types of experimental design that can yield clustered data. Here, we provide a statistical model for intra cluster correlation and systematically investigate a range of methods for analyzing clustered data.

For cluster data, mixed-effects regression models (MRMs) usually include random subject effects to account for the similarity among repeated measures for a given subject. The variance of the random subject effects, which represents between-subjects (BS) variation and the error variance, which represents within-subjects (WS) variation, are usually considered to be homogeneous across subject groups or levels of covariates. However, in reality, the homogeneous variance assumption, both within- and between-subjects, can be violated, and the random subject effects can further be correlated with the error terms.

If one analyzes clustered data with a models that assumes independence of observations, such as the generalized linear models, and ignores the clustering of observations, while the regression coefficients will be estimated without bias, the standard errors of time-invariant predictors will be underestimated see, e.g., Maas and Hox, 2005 Maas, C. J., Hox, J. J. (2005).

Multilevel models (MLMs) are very popular methods to analyze clustered or hierarchical structure. It is also used to analyze repeated measures data. Frequently in educational psychology research, observations have a hierarchical structure (Raudenbush and Bryk 2002). Students are nested within classrooms; children are nested within families, or teachers are nested within schools. MLMs are conceptually similar to multiple regression in that an outcome variable is linearly predicted from multiple independent variables (a.k.a. covariates or predictors). An assumption of traditional regression models such as ordinary least squares for continuous outcomes or logistic regression for binary outcomes is that observations (or the residuals) are independent. But cluster data violated this independence assumption, because observations are correlated within cluster. MLMs relax the independence assumption by directly modeling the clustered structure of the data through the inclusion of random effects or by directly modeling an alternative covariance structure for the residuals that accounts for the dependence of observations. Many different

types of parameters are present within MLMs including fixed effects at each level, random effects, their variance, and possible covariance components, as well as standard error estimates of each parameter.

Often, data have a clustered (panel or tabular) structure, that means observations are correlated within cluster. But classical statistics assumes that observations are independent and identically distributed (iid). Applied independence assumption to clustered data, may lead to false results. In contrast, clustered data assumes two sources of variation, within cluster and between clusters. Random effect models is crucial for modeling cluster data (two and three level modeling), which allow each cluster have their own random intercept and random slope respectively. Also, random effect models allow for the possibility that one or more covariates have effects that vary from unit (cluster, subject) to unit. In statistics the random component can be estimated via various approaches such as maximum likelihood estimation, restricted maximum likelihood estimation.

Estimating the random component from the random effect models using various statistical tool like ML (maximum likelihood), REML (restricted maximum likelihood), AGQ (adaptive Gaussian quadrature), PQL (penalized quasi likelihood) is varied with no of cluster and cluster size for continuous and discrete outcome, which would be main concerned in this thesis. Many of the concerns with continuous outcomes are magnified with binary outcomes. When the outcome variable was binary (i.e., requiring the use of multilevel logistic regression), in to small numbers of clusters being problematic, cluster size played a much larger role with binary outcomes than with continuous outcomes. Even with a very large number of clusters, conditions with a cluster size of five or fewer produced estimates that exhibited much more bias than their counterparts with continuous outcomes (Austin, P.C 2010; Clarke 2008; Moineddin et al. 2007). Another difficulty encountered with binary outcomes is that estimation of the models is less straightforward

As implemented in statistical packages, linear mixed models assume that we have modeled the dependency structure correctly the random effects and within-unit residual errors follow normal distributions, and that these have constant variance. While it is possible to some extent to check these assumptions through various diagnostics, a natural concern is

that if one or more assumptions do not hold, our inferences may be invalid. Fortunately, it turns out that linear mixed models are robust to violations of some of their assumptions.

To outline the contents of this paper, first, a brief overview of cluster data and the multilevel models (two and three level models). Then, the MLMs with a small number of clusters is reviewed in detail. Third, an illustrative simulation study is provided that depicts the potential issues in modeling data with a small number of clusters with MLMs. Fourth, compare the different methods to estimate parameter with increase the cluster and cluster size.

An important problem in multilevel modeling is what constitutes a sufficient cluster and cluster size for accurate estimation. In multilevel analysis, the major restriction is often the higher-level cluster size. In this thesis, a simulation study is used to determine the influence of different cluster sizes at the group level on the accuracy of the estimates (regression coefficients and variances). In addition, the influence of other factors, such as the lowest-level cluster size and different variance distributions between the levels (different intraclass correlations), is examined. The results show that only a small cluster size at level two and three leads to biased estimates. The bias estimates of multilevel models, which differ in terms of the number of levels (e.g., 2, 3), scale of the outcome variable (e.g., continuous, categorical). To estimate MLMs without bias, adequate cluster and cluster sizes must be obtained, since MLMs are often estimated with maximum likelihood (ML) and restricted maximum likelihood methods. ML and REML estimates are asymptotically unbiased, meaning that they perform well as large sample sizes approach but are known to behave less desirably with smaller samples sizes, particularly when the number of clusters is small.

It is important that researchers are aware of the MLMs when sample size is small. Methods in a simulated analysis of cluster data, the impact of a cluster size was examined on both a continuous and dichotomous outcome in a simple two and three level multilevel models. For different linear and logistic models were examined: models with a group-level covariate; models with an individual-level covariate and models with an aggregated group-level covariate. The study evaluated further whether the impact of small cluster size differed depending on the total number of clusters. When the number of clusters were

large, neither fixed nor random components were affected by small cluster size. As the number of clusters decreased, the estimates of random effects were inflated. Furthermore, group-level variance estimates were more affected by random components.

Although a specific sample size to ensure unbiased estimates cannot be pinpointed, a few guidelines have been suggested such as 30 clusters (also deemed macrounit, level-2 sample size, site, or individuals for clustered data) with a cluster size of 30 (microunit, level-1 sample size, or clustered data) in Kreft (1996), a minimum of 20 clusters (Snijders and Bosker 2012), or 50 clusters with a cluster size of 20 for cross-level interactions or 100 clusters with 10 units each if the main interest is in the variance components (Hox 1998, 2010). From a design perspective, Snijders and Bosker (1993) also advise against MLMs if the number of clusters is below 10. However, in applied settings, the demands of these recommendations are not always realized, leading to potentially biased results. For instance, in a review by Dedrick et al. (2009), using the 30/30 guideline, of the 99 studies reviewed using MLMs between 1999 and 2003 in 13 journals from education, psychology, and sociology, 21 % had sample sizes that would not meet the recommendation.

For a general introduction to multilevel modeling of hierarchical or cluster data, we refer to Snijders and Bosker (1999), Heck and Thomas (2000), Raudenbush and Bryk (2002), and Hox (2002). The maximum likelihood (ML) and restricted maximum likelihood (REML) estimation methods used commonly in multilevel analysis are asymptotic, which assumption that the sample size must be sufficiently large. This raises questions about the acceptable lower limit to the sample size, and the accuracy of the estimates and the associated with relatively small cluster sizes. In multilevel studies, the main problem is usually the sample size at the group level, because the group-level sample size is always smaller than the individual-level sample size.

Currently there are few sample size guidelines referenced in the literature. One the rule of the thumb proposed for designs in which individuals are nested within groups calls for a minimum of 30 units at each level of the analysis. This rule of thumb is commonly cited (Hox, 1998; Mass & Hox,2004) and was further developed by Hox (1998) who recommended a minimum of 20 observations (level-1) for 50 groups (level-2). Maas and Hox (2004,2005) who examined conditions in which the number of level-2 units ranged 30

to 100 and the number of level-1 units per level-2 unit ranged from 5 to 50, found the less bias in the variance component estimate, but still reported substantial in making inference about the variance components when the number of level-2 units only 30.

2. Overview of Two-Level Models

To illustrate the MLM, consider the following two-level models for continuous and discrete outcomes as

$$\text{Level -1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + r_{ij}$$

$$\text{Level-2: } \beta_{0j} = \gamma_{00} + \gamma_{01}W_{1j} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_{1j} + u_{1j}$$

where Y_{ij} is a response variable for the i^{th} observation in the j^{th} cluster, X_{1ij} is the value of the level-1 predictor for the i^{th} observation in the j^{th} cluster, r_{ij} is the level-1 residual for the i^{th} observation in the j^{th} cluster, γ are fixed effect estimates, W_{1j} is the value of the level-2 predictor for the j^{th} cluster, u_{0j} and u_{1j} are the random effects of the j^{th} clusters are considered to be a sample from the larger population. For random effect models the parameter β_0 and β_1 follows the multivariate normal distribution with mean μ and variance-covariance matrix Σ and the error term also follow normal distribution with mean 0 and variance σ^2 .

3. Overview of Three Level Models

Many of the concepts for two-level models translate to three-level models. Therefore, there is a need for developing methods for a three-level analysis with random intercept and random slope models which can be written as

$$\text{Level-1: } Y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{1ijk} + r_{ijk}$$

$$\text{Level-2: } \beta_{0jk} = \gamma_{00k} + \gamma_{01k}W_{1jk} + u_{0jk}$$

$$\beta_{1jk} = \gamma_{10k} + \gamma_{11k}W_{1jk} + u_{1jk}$$

$$\text{Level-3: } \gamma_{00k} = \delta_{000} + \delta_{001}Z_{01k} + v_{00k}$$

$$\gamma_{10k} = \delta_{100} + \delta_{101}Z_{01k} + v_{10k}$$

Where Y_{ijk} is the response variable for i^{th} students of the j^{th} classroom and k^{th} school? X_{1ijk} is the value of the level-1 predictor for the i^{th} students in the j^{th} classroom and k^{th}

school, r_{ijk} is the level-1 residual for the i^{th} students in the j^{th} classroom and k^{th} schools and γ are fixed effect estimates, W_{1jk} is the value of the level-2 predictor for the j^{th} classroom and k^{th} school, u_{0jk} and u_{1jk} are the random effects of the j^{th} classroom and k^{th} school. Z_{01k} is the level-3 predictor for the k^{th} school. Here for random effect models the parameter follows multivariate normal distribution with mean μ and variance-covariance matrix Σ and the error term follows the normal distribution with mean 0 and variance σ^2 .

4. Methodology

In this paper we illustrate the two and three level models for discrete and continuous outcome. Different methods are applying for estimating the parameter for continuous and discrete outcome. There are many methods for estimating parameter in continuous outcome like as Maximum likelihood Estimation (MLE), Restricted Maximum Likelihood Estimation. Another different method applies for estimating discrete outcome like as Restricted Maximum Likelihood Estimation (MLE), Penalized Quasi Likelihood (PQL), and Adaptive Gaussian Quadrature (AGQ). In two and three level models we estimate relative bias and mean square error for measure the accuracy in different methods. Comparing different methods based on the criteria of mean square error and relative bias. In multilevel modeling increase the number of cluster and cluster size relative bias and mean square error decrease different methods for discrete and continuous outcome in two and three level models. In this paper we estimate the relative bias and mean square error for random intercept parameter for two and three level models and compare different methods in discrete and continuous outcome.

5. Result and Discussion

In order to identify the effect of varying the no. of cluster and cluster size on estimated parameter using the different estimation technique, we design this simulation part. We simulated data from random effect models of two level and three level, considering the continuous and discrete parameter respectively and different estimation tools are used, such as, maximum likelihood estimation and restricted maximum likelihood estimation (for

continuous outcome), Penalized Quasi likelihood and Adaptive Gaussian quadrature (for discrete outcome).

Simulation Study for two level model

Two levels: Continuous Outcomes (Cluster no. fixed)

Here we also display the relative bias and mean square error for continuous outcome. We compare REML and ML method

Figure 1: Relative bias of random intercept and random slope for continuous outcome considering cluster no. fixed

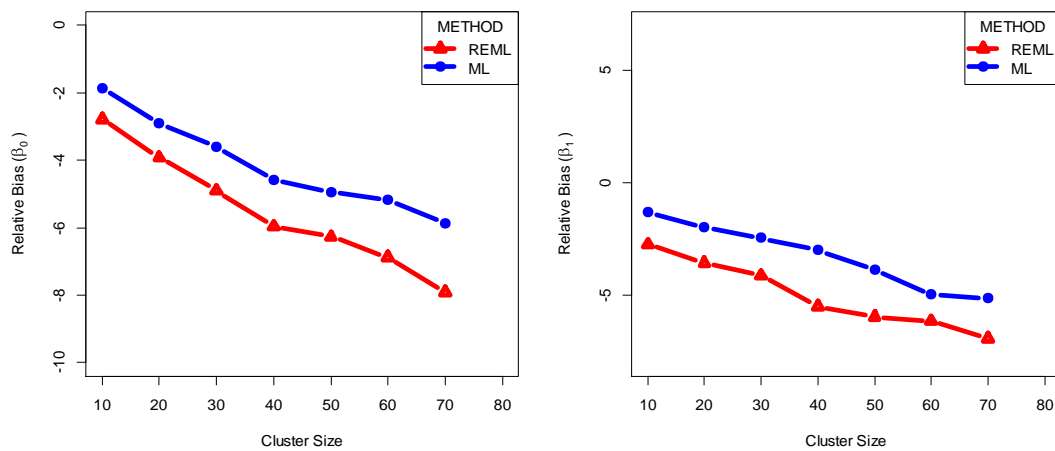
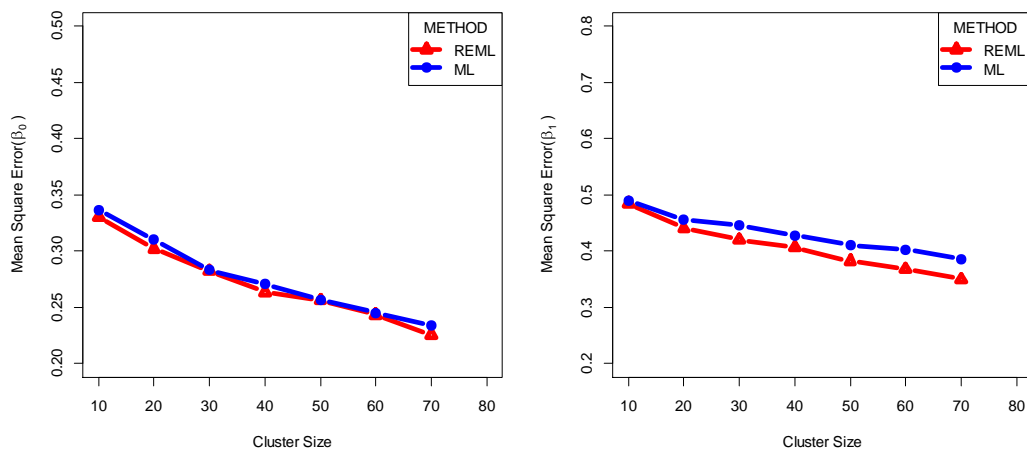


Figure 2: Mean square error of random intercept and random slope for continuous outcome considering cluster no. fixed



In the above graph 1 display the result for relative bias for random intercept and random slope parameter when the number of clusters is fixed. Here observe that the relative bias for random intercept and random slope parameter is decreasing with increasing the cluster size for both methods but in REML methods relative bias more decreasing than the ML methods.

Also, in the figure 2 graphical representation of mean square error for random intercept and random slope parameter for fixed cluster. From the graph 2 we can see that when the cluster size is varying then the mean square error is decreasing for random intercept and random slope parameter in two level modals.

Two levels: Continuous Outcomes (Cluster size fixed):

Figure 3: Relative bias of random intercept and random slope for continuous outcome considering cluster size fixed

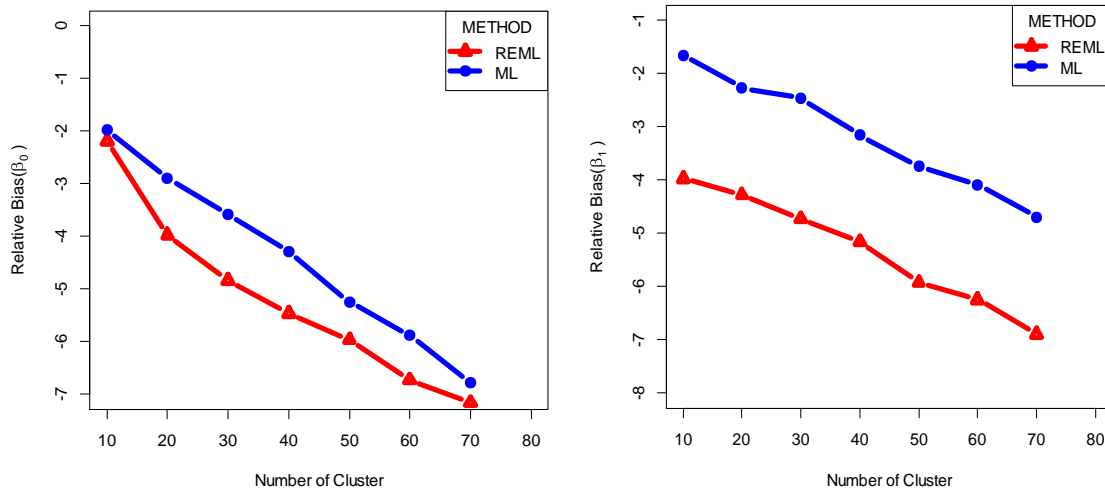
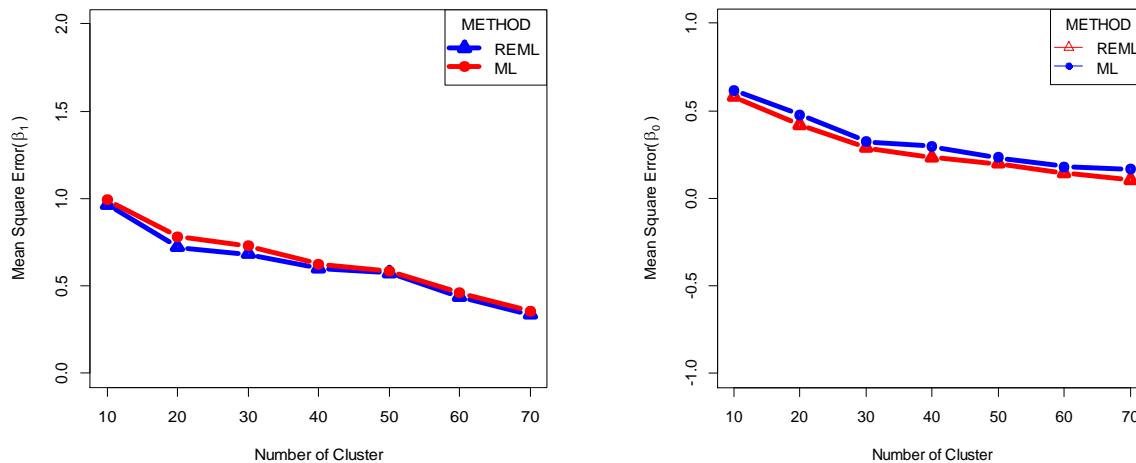


Figure 4: Mean square error of random intercept and random slope for continuous outcome considering cluster size fixed



From the figure 3 we observe when cluster size fixed but number of cluster varying then relative bias for random intercept and random slope parameter are decreasing in ML and REML methods but REML methods give less relative bias.

Figure 4 show the mean square error for random intercept and random slope parameter decreasing with increasing the number of cluster when the cluster size fixed. Here we observe that in both methods mean square error always decreasing with increasing the number of cluster but in REML mean square error more decreasing than ML methods.

Coverage which means that how many times the true parameter values are falls in the desire confidence interval. If we find the coverage for 95% confidence interval then it indicates that how many times the true parameter values are falls in 95% confidence interval. Coverage can often be compared to the nominal confidence of the interval. Low coverage meant that biased estimates or anticonservative standard error estimates or both. High coverage probabilities indicate overly conservative standard error estimates. I would like to find the coverage in random intercept and random slope parameter.

Table 1: Coverage of random intercept and random slope for continuous by REML methods

Parameter	No. of cluster	Cluster size	Coverage rate (%)
Intercept (β_0)	10	10	91.6
	20		92.8
	60		93.4
	80		94.4
	120		95.6
Slope (β_1)	10	10	91.6
	20		92.2
	60		93.3
	80		94.1
	120		95.4

Tables 1 display the coverage rate increasing with increasing the number of cluster when cluster size fixed. Initially when number of cluster 10 and cluster size fixed 10 then the coverage rate is 91.6% for random intercept and random slope parameter. But when we increase the number of cluster then the coverage rate also increases. Lastly when the number of clusters is 120 and cluster size 10 then coverage rate is for $\beta_0 = 95.6\%$ and $\beta_1 = 95.4\%$ which is equal to 95% confidence intervals.

Two levels: Discrete Outcomes (Cluster no. fixed)

Since in the above section we justify ML and REML methods only for continuous outcome, now we also justify this result in discrete outcome variables. Her we use three methods for estimating the parameter in two level models and we only consider the random intercept parameter for discrete outcome. Her we use three methods for estimating parameter

- 1) REML (Restricted maximum likelihood)
- 2) PQL (Penalized quasi likelihood)
- 3) AGQ (Adaptive Gaussian quadrature)

Here graphically shows the result for discrete outcome for random intercept parameter

Figure 5: Relative bias of random intercept for discrete outcome considering cluster size fixed

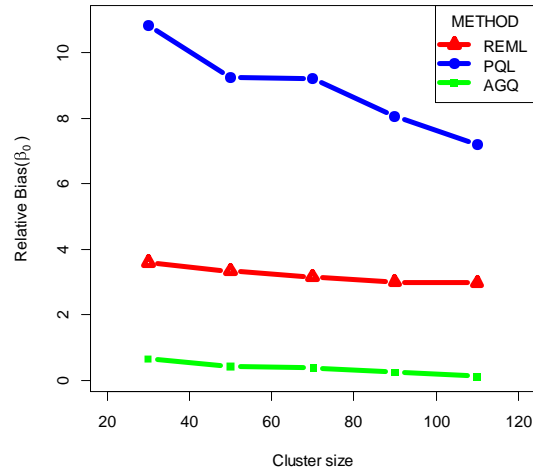


Figure 6 Mean square error of random intercept for discrete outcome considering cluster no. fixed

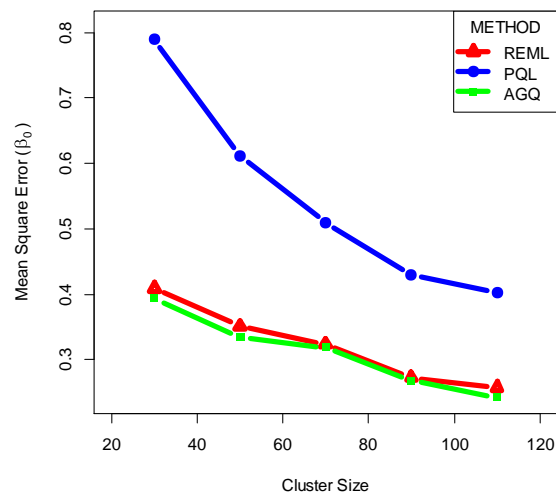


Figure 5 the graphical representation of relative for random intercept parameter for discrete outcome. Here we see that relative bias for random intercept parameter decreasing with increasing the cluster size. In figure 6 we compare three methods and, in each method, relative bias decreasing when we increasing the cluster size. Graph (figure 6) display the

result mean square error for random intercept parameter when number of clusters fixed. In this figure mean square error in random intercept parameter decreasing with increases the cluster size.

Two levels: Discrete Outcomes (Fixed cluster size)

Figure 7: Relative bias of random intercept for discrete outcome considering cluster size fixed

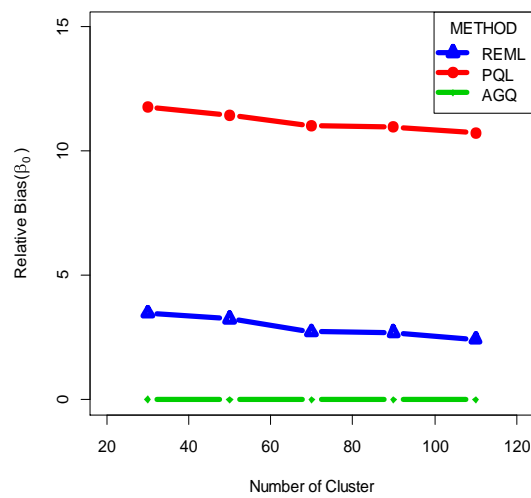
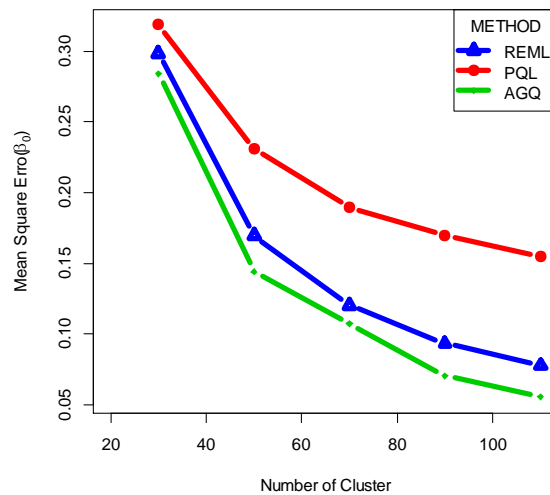


Figure 8: Mean square error of random intercept for discrete outcome considering cluster size fixed



The graph (figure 7) shows after increasing the number of cluster relative bias also decreasing three methods. From figure 8 mean square error also decrease with increase the number of clusters. Since adaptive Gaussian quadrature methods is better among the three methods. We find the coverage rate for discrete outcome. Here the table for coverage rate is given below

Table 2: Coverage of random intercept for discrete outcome by AGQ methods

Parameter	Number of cluster	Cluster size	Coverage rate (%)
Intercept (β_0)	5	15	68.4
	10		82.6
	20		89.2
	60		93.8
	80		94.0
	120		95.2

Table 2 show the coverage rate for 95% confidence interval for different number of clusters. When cluster size fixed 15 and number of cluster 5 then the coverage rate for random intercept parameter is 68.4%, when we estimate the parameter by adaptive Gaussian quadrature methods. But when increase the number of cluster 120 then the coverage rate also increases in 95.2% which means that with increase the number of clusters the accuracy of adaptive Gaussian quadrature methods closed to the true parameter value.

Simulation Study for three level models

Three levels: Continuous Outcome (varying level-2 cluster no.):

Figure 9: Relative bias of random intercept for continuous outcome considering (level-2 cluster no.) fixed

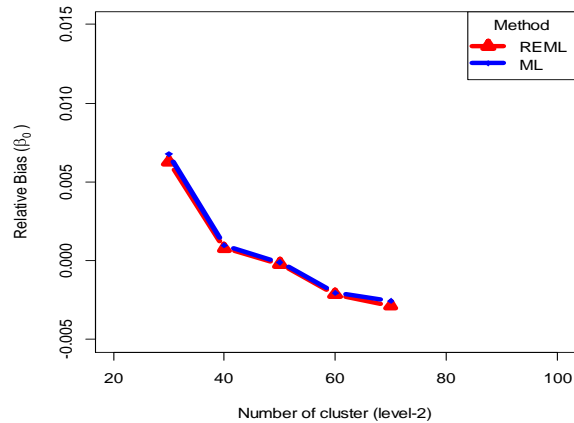


Figure 10: Mean square error of random intercept for continuous outcome considering (level-2 cluster no.) fixed

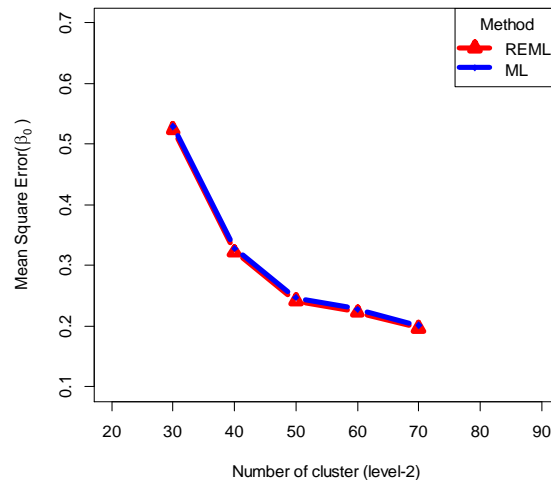


Figure 9 display the relative bias in continuous outcome for level-2 parameter when number of cluster (leve-3) fixed. From the graph we observe that with increase the number of cluster (level-2) relative bias decreasing in three level models. In both methods relative

bias decreasing but REML methods produce less relative bias in three level models when we try estimate relative bias for (level-2) in three level models. Also, here from the figure 10 the mean square error in level-2 parameter (level-2) shows that with increasing the number of classrooms mean error also decreasing for three level models. Again, we estimate the coverage rate for level-2 parameter and check the parameter how many times fall the true parameter value in 95% confidence interval. Table 3 shows the coverage rate for level-2 parameter in three level models

Table 3: Coverage of random intercept (level-2) for discrete outcome by REML methods

Parameter	No.of cluster (level-2)	No.of cluster (level-3)	Coverage rate (%)
Intercept (β_0)	20	15	46.2
	40		45.0
	60		47.4
	80		45.2
	120		49.0

Table 3 provides the coverage rate for level-2 parameter. Here Initially when number of cluster (level-2) is 20 then coverage rate only 46.2 which is so far from the 95% confidence interval. Although we increase the number of classrooms in 120 then the REML methods provided low coverage rate and it is so far from 95% confidence interval.

Three levels: Continuous Outcome (varying level-3 cluster no.)

Figure 11: Relative bias of random intercept for continuous outcome considering (level-3 cluster no.) fixed

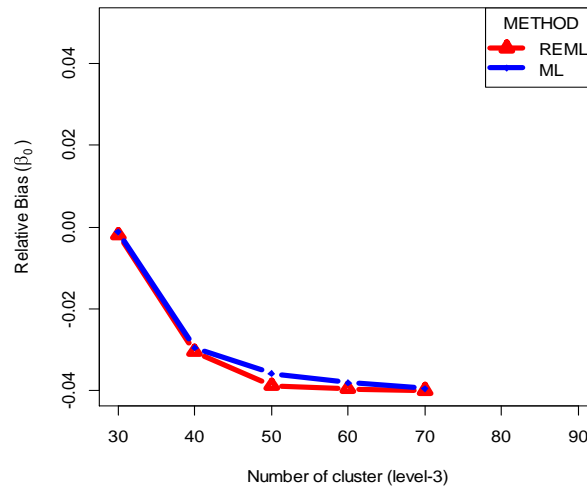
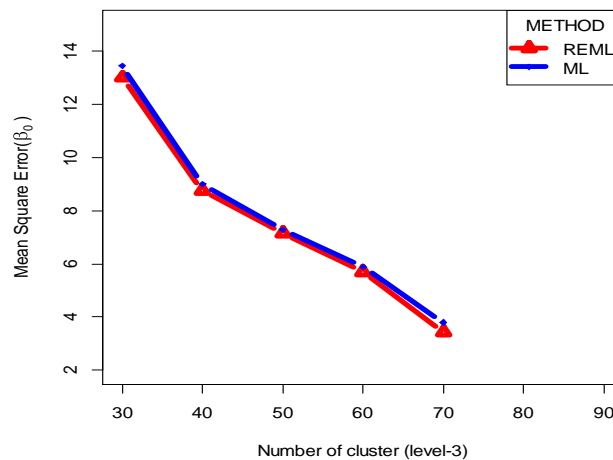


Figure 12: Mean square error of random intercept for continuous outcome considering (level-3 cluster no.) fixed



The graph (figure 11) displays the relative bias in level-3 parameter always decrease in this both methods but REML methods produce less relative bias. Based on the criterion of mean square error we compare ML and REML methods for level-3 parameter. From figure 12 shows that mean square error decreasing with increasing the number of cluster (level-3). Now we also find the coverage rate for level-3 parameter in three level models.

Table 4: Coverage of random intercept (level-3) for continuous outcome by REML methods

Parameter	No. of cluster (level-3)	No. of cluster (level-2)	Coverage rate (%)
Intercept (β_0)	20	15	90.4
	40		93.4
	60		93.6
	80		94.6
	120		95.3

Table 4 provide coverage rate for different number of schools in three level models for 95% confidence interval. When number of cluster (level-3) is 20 then the coverage rate is 90.4%. After increase number of cluster (level-3) the coverage rate also increases which is approximate to 95%. But three levels when we estimate coverage rate for level-2 parameter then it provides low coverage rate. So, we conclude that in three level when estimate coverage rate for level-3 parameter, it is provided better coverage rate and close to desire confidence interval but level-2 parameter provide low coverage rate.

Three levels: Continuous Outcome (varying level-2 and level-3 cluster no. simultaneously):

Figure 13: Relative bias of random intercept for continuous outcome varying level-2 and level-3 cluster no. simultaneously

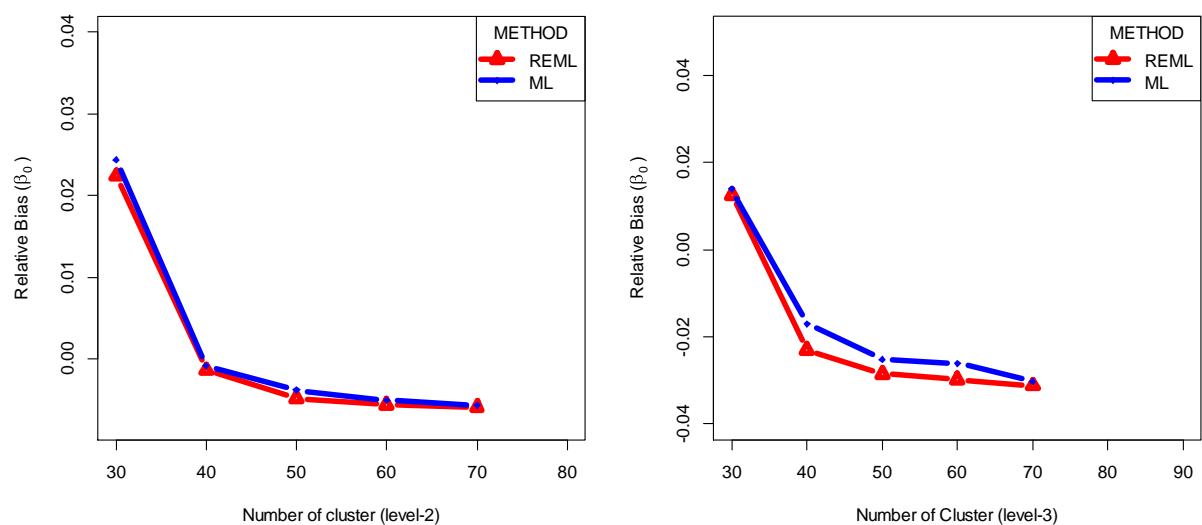
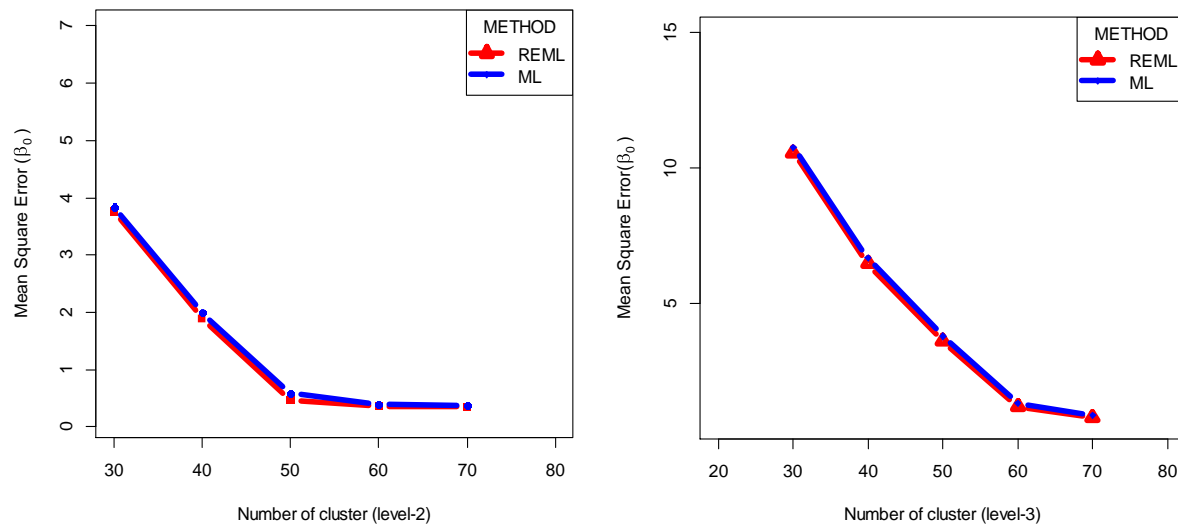


Figure 14 Mean square error of random intercept for continuous outcome varying level-2 and level-3 cluster no. simultaneously



The figure 13 shows that relative bias and for random intercept parameter when varying level-2 and level-3 together. In this both methods for level-2 and level-3 we only consider the random intercept parameter and compare the result in ML and REML methods.

The graph 14 provide the mean square error in different number of cluster (level-2) and different number of cluster (level-3). Here also mean square error decreasing with increasing the number of cluster (level-2) and number of cluster (level-3) simultaneously. Now we estimate the coverage rate for varying level-2 and level-3 simultaneously.

Table 5: Coverage of random intercept (level-2 and level-3) for continuous outcome by REML methods

Parameter	No.of Cluster (level-2)	No.of Cluster (level-3)	Coverage rate(%) (level-2)	Coverage rate (%) (level-3)
Intercept (β_0)	5	20	33.0	90.4
	20	40	26.0	93.6
	30	60	20.4	65.0
	40	80	24.0	95.6
	60	120	17.2	93.2

Table 5 provide the coverage rate in level-3 parameter. When number of cluster (level-2) is 5 and number of cluster (level-3) is 20 then the coverage rate for level-2 and level-3 parameter is 33% and 90.4% receptively. Here level-2 parameter provides the low coverage rate but level -3 parameter provide better coverage rate.

Three levels: Discrete Outcomes (varying level-2 cluster no.)

Figure 15: Relative bias of random intercept for discrete outcome varying (level-2 cluster no.)

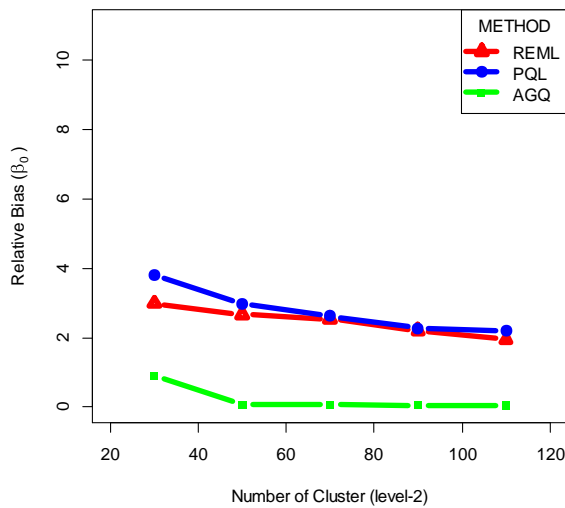
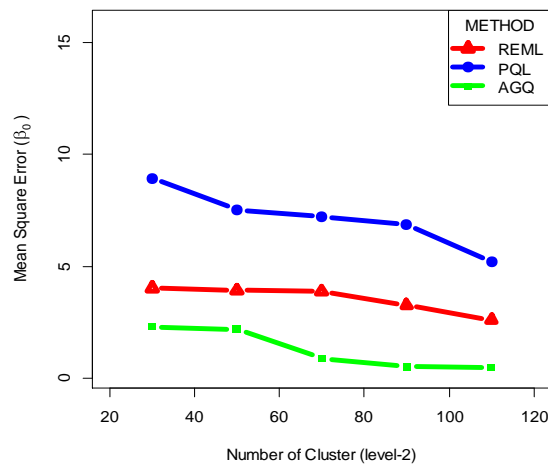


Figure 16: Mean square error of random intercept for discrete outcome varying (level-2 cluster no.)



From the graph 15 provide that the relative bias in level-2 parameter decrease when increase the number of cluster (level-2). Here relative bias in level-2 parameter by AGQ methods provided less relative bias in three level models.

The (figure 16) we observe that mean square error decreasing in three methods when increasing the number of cluster (level-2) in three level models for discrete outcome. Now we also estimate coverage rate for level-2 parameter when estimate parameter by AGQ methods.

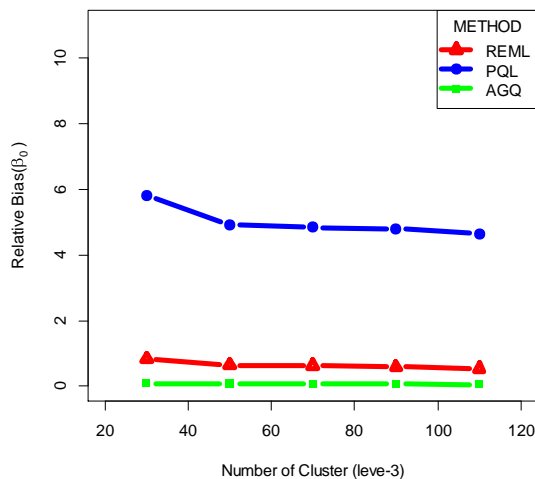
Table 6: Coverage of random intercept (level-2) for discrete outcome by AGQ methods

Parameter	No.of cluster (leve-2)	No.of cluster (leve-3)	Coverage rate (%)
Intercept (β_0)	20	15	47.1
	40		53.2
	60		43.6
	80		36.8
	100		51.0
	120		50.5

In the above table 6 which shows the coverage rate in level-2 parameter for different number of cluster (level-2) when number of cluster (level-3) fixed. Here we observe that when we increase the number of cluster (level-2) then it provides the low coverage rate in level-2 parameter which is so far from the 95% confidence interval.

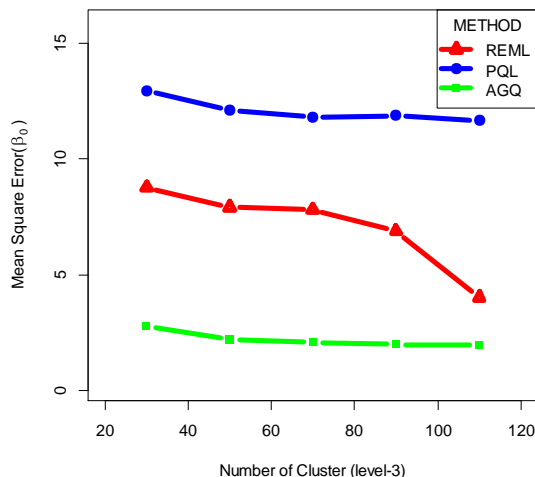
Three levels: Discrete Outcomes (varying level-3 cluster no.)

Figure 17: Relative bias of random intercept for discrete outcome varying (level-3 cluster no.)



The (figure 17) provide the relative bias in level-3 parameter for discrete outcome in three level models. In this graph relative bias decreasing with increasing the number of cluster (level-3) when number of cluster (level-2) is fixed. Here adaptive Gaussian quadrature methods provided less relative bias among the three methods

Figure 18: Mean square error of random intercept for discrete outcome varying (level-3 cluster no.)



In the above figure 18 which displayed the mean square error for level-3 parameter when number of cluster (level-2) fixed. With increasing the number of cluster (level-3) mean square error is decreasing and AGQ methods provided less mean square error other than two methods. Now we estimate coverage rate for level-3 parameter when estimate level-3 parameter by AGQ methods.

Table 7: Coverage of random intercept (level-3) for discrete outcome by AGQ methods

Parameter	No.of cluster (level-3)	No.of cluster (level-2)	Coverage rate (%)
Intercept (β_0)	20	15	88.0
	40		87.0
	60		89.2
	80		91.5
	120		93.0

In the above table 37 when we estimate parameter by AGQ then the coverage rate is 88% in level-3 parameter for number of clusters 20. When increase the number of clusters we observe from the table the coverage rate for in level-3 parameter also increases and close to the 95% confidence interval.

6. Conclusion

In this thesis paper the main idea is to effect of sample size in multilevel models for cluster data when we estimate random effect models parameter. There are several authors have try to estimate fixed effect parameter in multilevel models for continuous and discrete outcome variable. They also try to estimate parameter for different methods such as restricted maximum likelihood (REML), maximum likelihood (ML) for continuous outcome and adaptive Gaussian quadrature (AGQ), penalized quasi likelihood for discrete outcome. Many authors estimate the multilevel model's parameter and compare the different methods. Given the potential problems of small sample sizes, several simulation studies have been designed to examine the effect of small sample sizes, at different levels of analysis, on various multilevel results (e.g., variance estimates, fixed effects estimates).

In the above all discussion, we can say that the researcher more conscious about sample size because there is more impact when estimates two and three level models parameter. It is important that researchers are aware of and acknowledge the potential untrustworthiness present with MLMs when sample size is small. To estimate multilevel models without bias, adequate sample sizes must be obtained, since multilevel are often estimated with different methods for discrete and continuous outcome and all methods affect by small sample size.

Finally, we can say that REML method provided better result of continuous outcome for two and three level model and AGQ method provided less relative bias and mean square error for two and three level model.

REFERENCES

- [1] P.C. Austin, Estimating multilevel logistic regression models when the number of clusters is low: a comparison of different statistical software procedures, *Int. J. Biostat.* 6 (2010) 16. <https://doi.org/10.2202/1557-4679.1195>.
- [2] J. Cohen, Determining sample sizes for surveys with data analyzed by hierarchical linear models, *J. Official Stat.* 14 (1998) 267–275.
- [3] P. Clarke, When can group level clustering be ignored? Multilevel models versus single level models with sparse data, *J. Epidemiol. Community Health*, 62 (2008) 752–758. <https://doi.org/10.1136/jech.2007.060798>.
- [4] R.F. Dedrick, J.M. Ferron, M.R. Hess, K.Y. Hogarty, J.D. Kromrey, T.R. Lang, J.D. Niles, R.S. Lee, Multilevel modeling: a review of methodological issues and applications, *Rev. Educ. Res.* 79 (2009) 69–102. <https://doi.org/10.3102/0034654308325581>.
- [5] H. Goldstein, W. Browne, J. Rasbash, Partitioning variation in multilevel models, *Understand. Stat.* 1 (2002) 223–231. https://doi.org/10.1207/S15328031US0104_02.
- [6] J. Hox, Multilevel modeling: when and why, in: I. Balderjahn, R. Mathar, M. Schader (Eds.), *Classification, Data Analysis, and Data Highways*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1998: pp. 147–154. https://doi.org/10.1007/978-3-642-72087-1_17.
- [7] J. Hox, *Multilevel analyses: techniques and applications* (2nd ed.). Mahwah, NJ: Erlbaum. (2010). <https://doi.org/10.4324/9781315650982>.
- [8] I.G.G. Kreft, Are multilevel techniques necessary? An overview, including simulation studies. Unpublished manuscript, California State University, Los Angeles. (1996).

- [9] C.J.M. Maas, J.J. Hox, Robustness issues in multilevel regression analysis, *Statistica Neerland.* 58 (2004) 127–137. <https://doi.org/10.1046/j.0039-0402.2003.00252.x>.
- [10] C.J.M. Maas, J.J. Hox, Sufficient sample sizes for multilevel modeling, *Methodology.* 1 (2005) 86–92. <https://doi.org/10.1027/1614-2241.1.3.86>.
- [11] O. Paccagnella, Sample size and accuracy of estimates in multilevel models: new simulation results, *Methodology.* 7 (2011) 111–120. <https://doi.org/10.1027/1614-2241/a000029>.
- [12] D.M. McNeish, Modeling sparsely clustered data: Design-based, model-based, and single-level methods, *Psychol. Methods.* 19 (2014) 552–563. <https://doi.org/10.1037/met0000024>.
- [13] R. Moineddin, F.I. Matheson, R.H. Glazier, A simulation study of sample size for multilevel logistic regression models, *BMC Med. Res. Methodol.* 7 (2007) 34. <https://doi.org/10.1186/1471-2288-7-34>.
- [14] S.W. Raudenbush, A.S. Bryk, *Hierarchical linear models: applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage. (2002). <https://psycnet.apa.org/record/1992-97494-000>
- [15] F.E. Satterthwaite, An approximate distribution of estimates of variance components, *Biometrics Bull.* 2 (1946) 110. <https://doi.org/10.2307/3002019>.
- [16] T.A.B. Snijders, R.J. Bosker, Standard errors and sample sizes for two-level research, *J. Educ. Stat.* 18 (1993) 237–259. <https://doi.org/10.3102/10769986018003237>.
- [17] T.A.B. Snijders, R.J. Bosker, *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage. (2012).