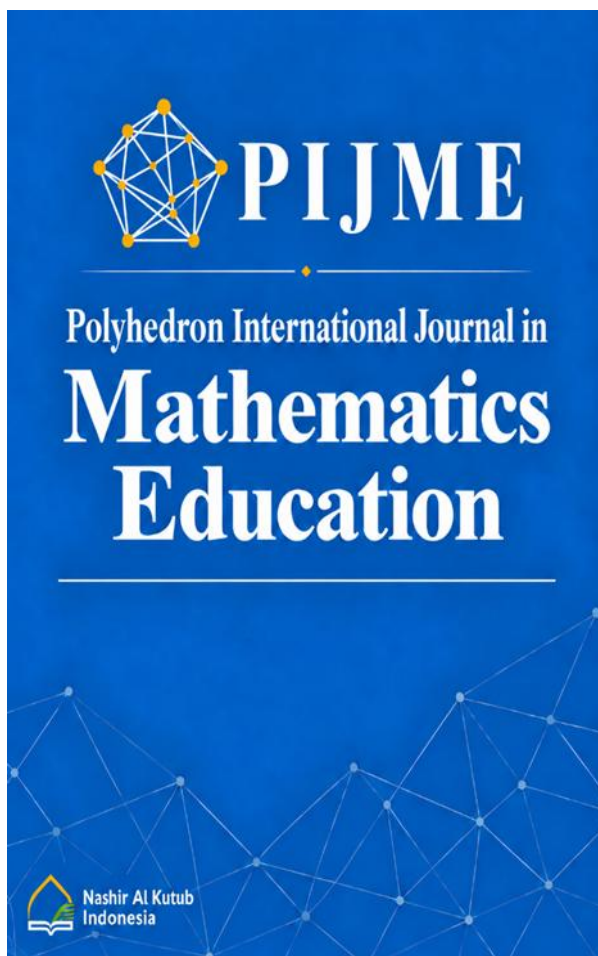




Bayesian hierarchical modeling of factors associated with students' educational status in less developed areas: A case study of Ghor Province, Afghanistan

Mohammad Nasim Wafa

Herat University, Herat, Afghanistan,
wangf701@nenu.edu.cn

To cite this article:

Wafa, M.N. (2026). Bayesian hierarchical modeling of factors associated with students' educational status in less developed areas: A case study of Ghor Province, Afghanistan. *Polyhedron International Journal in Mathematics Education*, 4(2), 193 – 210.

To link to this article:

<https://nakiscience.com/index.php/pijme>

Published by:

Nasir Al-Kutub Indonesia

Residential Street Kila Rengganis, Block I, Number 11, Labuapi, Indonesia, 83361

Bayesian hierarchical modeling of factors associated with students' educational status in less developed areas: A case study of Ghor Province, Afghanistan**Mohammad Nasim Wafa**Herat University, Herat, Afghanistan, wangf701@nenu.edu.cn**Abstract**

Students' educational status in remote, data-scarce regions is shaped by several interrelated factors, yet evidence that quantifies these associations together with their uncertainty is scarce. This study examined factors associated with students' educational status in less developed areas of Ghor Province, Afghanistan, using Bayesian regression and Bayesian hierarchical modeling. A cross-sectional correlational survey was conducted in 2019 with 372 students in Grades 11 and 12 from 32 schools in the Firozkoh, Dawlatyar, and Lal wa Sarjangal districts. A 35-item questionnaire measured economic status and family support, educational facilities and infrastructure, teacher quality and teaching process, geographical accessibility, environmental conditions, and perceived educational status. A single-level Bayesian regression and a Bayesian hierarchical model with school-specific random intercepts and district indicators were fitted, and inference was based on posterior distributions, 95% credible intervals, and Markov chain Monte Carlo diagnostics. All five factors were positively associated with educational status; teacher quality and teaching process showed the largest posterior mean coefficient (0.31; 95% credible interval [0.21, 0.41] in the single-level model), followed by educational facilities and infrastructure (0.26). The hierarchical model indicated between-school heterogeneity (posterior mean between-school standard deviation = 0.28) and showed better expected predictive performance than the single-level model (LOOIC lower by 36.8 points). Because the design was cross-sectional and the sample was a convenience sample, the findings indicate associations rather than causal effects. They nevertheless suggest that teacher quality, school infrastructure, and household, geographical, and environmental conditions deserve attention in educational planning for such regions.

Article History

Received:

August 1, 2026

Revised:

September 30, 2026

Accepted:

October 10, 2026

Published Online:

October 12, 2026

Keywords:

Bayesian hierarchical model;

Educational status;

Less developed areas;

Multilevel modeling

1. Introduction

Statistical modeling is an essential tool for analyzing complex phenomena and the relationships between variables across the sciences. In many applied problems, the researcher's goal is not only to determine whether variables are related, but also to estimate the strength of the associations between different factors and to quantify the uncertainty of the results. Classical regression methods, such as linear, logistic, and

multinomial regression, have played an important role in data analysis for decades. However, these methods rely on assumptions such as a sufficiently large sample, fixed parameters, and independence of observations, which may not be fully met in many practical problems, especially in less developed areas.

Bayesian inference offers an alternative framework. Model parameters are treated as random quantities, and uncertainty about them is expressed through probability distributions that combine prior information with the observed data to yield the posterior distribution of the parameters (van de Schoot et al., 2021). Instead of a single point estimate and a p-value, the analyst obtains the full posterior distribution of each coefficient, from which credible intervals and the probabilities of substantively meaningful statements can be derived directly (Kruschke & Liddell, 2018). Bayesian regression can be applied to data with limited sample size, complex structure, or incomplete information, and its posterior distributions are usually approximated with Markov chain Monte Carlo (MCMC) methods. Bayesian methods are not, however, an automatic remedy for small samples: with few observations or poorly chosen priors, the results can depend strongly on the prior specification (McNeish, 2016).

Education is a multidimensional phenomenon influenced by economic, social, environmental, and infrastructural factors. From a statistical perspective, analyzing it involves estimating the association of each explanatory factor with educational indicators while accounting for the other factors. Educational data are also typically hierarchical, with students nested in schools and schools located in regions, so single-level regression models may not adequately represent the variation in the data. Less developed regions add further difficulties: statistical data are limited, populations are small and dispersed, and study units are heterogeneous. In such settings, Bayesian hierarchical models can incorporate prior information, represent the clustered structure of the data, and express uncertainty in probabilistic terms. Their value lies mainly in the transparent quantification of uncertainty and in modeling the nested structure of the data, rather than in compensating for limitations of sampling or measurement.

Bayesian hierarchical models have been used increasingly in educational research. Fouskakis et al. (2014) proposed a Bayesian hierarchical model for the comparative evaluation of teaching quality indicators in higher education, in which posterior distributions combine prior information with observed data and provide a direct probabilistic interpretation of the indicators. Tomal et al. (2021) used a Bayesian hierarchical modeling approach to examine the impact of COVID-19 on students' marks, showing that such models can separate individual- and group-level sources of variation in longitudinal educational data. Mosia et al. (2025) analyzed multi-year data on student academic performance and examined the roles of mathematics competency, institutional context, and temporal variability within a Bayesian hierarchical

framework, and Mosia et al. (2026) applied Bayesian hierarchical modelling to multi-cohort data on the effects of academic orientation and advising on student retention and progression. Sun (2026) applied hierarchical Bayesian analysis to responses from 4,151 students on 110 prerequisite mathematical skills to identify struggling students and the topics that require additional instructional support.

Related applications extend beyond student performance. Gómez-Méndez and Amornbunchornvej (2024) used Bayesian hierarchical models to study income, education, and other poverty-related variables, and Pardo et al. (2025) applied Bayesian multilevel spatial models to uncover latent territorial structure in performance on the Colombian ICFES Saber 11 examination. Le et al. (2025) introduced and evaluated Bayesian multilevel compositional data analysis, and Montano and Reyes (2026) proposed a generalized Bayesian hierarchical framework for course-level learning evaluation. Studies of school-level determinants, such as Tilbe and Biru (2026), have highlighted the role of school, teacher, and leadership factors in students' academic performance, and reviews of educational interventions in developing countries point to the central role of pedagogy and teachers, although their conclusions vary with the studies included (Evans & Popova, 2016). In the Afghan context, Wafa et al. (2020) evaluated students' mathematical ability in Afghan schools using cognitive diagnosis models, and related methodological work has examined Bayesian t-Lasso regression for variable selection (Wafa & Shuja, 2024) and regression models based on the Birnbaum–Saunders distribution (Wafa et al., 2023).

Taken together, these studies show that Bayesian models can quantify uncertainty, perform reasonably in moderate samples when priors are chosen carefully, and separate individual-, school-, and regional-level sources of variation. However, most educational applications focus on student performance, teaching quality, or the prediction of academic success in relatively data-rich settings. Few studies have used Bayesian approaches to examine factors associated with students' educational status in less developed regions, and this gap is particularly evident for Afghanistan, where data are scarce, regional differences are large, and uncertainty is high. Ghor Province, with its mountainous terrain, dispersed settlements, and limited educational infrastructure, is a case in point.

Accordingly, the main objective of this study is to estimate, within a Bayesian framework, the associations between five contextual factors and students' educational status in less developed areas of Ghor Province, while accounting for the clustering of students within schools. Specifically, the study addresses three research questions: (RQ1) How are economic status and family support, educational facilities and infrastructure, teacher quality and teaching process, geographical accessibility, and environmental conditions associated with students' perceived educational status after mutual adjustment? (RQ2) How much do schools differ in students' educational status after these factors are taken into account? (RQ3) Does a Bayesian hierarchical model

that accounts for school clustering provide better expected predictive performance than a single-level Bayesian regression model? The study contributes context-specific empirical evidence from Ghor, a remote and data-scarce province for which quantitative evidence on schooling conditions is very limited, and shows how posterior distributions and credible intervals can be used to communicate uncertainty in such settings.

2. Methods

2.1 Research Design and Participants

This study used a quantitative, cross-sectional correlational design based on probabilistic modeling. The population comprised students in Grades 11 and 12 attending schools in Firozkoh, Dawlatyar, and Lal wa Sarjangal, Ghor Province. In total, 372 students from 32 schools completed the questionnaire. Convenience sampling was used under the prevailing field and accessibility conditions, and enrolled students who were willing to participate were included. The district subsamples comprised 172, 86, and 114 students, respectively. Across the 32 schools, the average number of participating students was 11.63. School selection procedures, district-specific school counts, and the range and median of school sample sizes require confirmation from the sampling records.

The data have a nested structure in which students (level 1) are nested within schools (level 2), and the schools are located in three districts. Because three districts are too few to estimate a district-level variance component reliably, district was not modeled as a third random level; instead, it was included as a school-level fixed effect with two indicator variables (Section 2.3).

2.2 Instrument

Data were collected using a researcher-made questionnaire organized around the study objectives and six constructs reflecting family resources, school facilities, teaching, accessibility, environmental conditions, and perceived educational status in Ghor Province. The construct definitions and item counts are described below. Internal consistency is reported in Table 3. The available manuscript does not document the administration language, translation or adaptation procedures, expert review, pilot sample, or factor-analytic results; consequently, content and construct validity cannot be established from the reported reliability coefficients alone. These details should be verified against the original instrument-development records.

The questionnaire contained 35 items rated on a five-point Likert scale coded from 1 (very little) to 5 (very much) and grouped into six constructs. Economic status and family support (six items) covered the family's economic status, its ability to meet educational needs, parental education, family support for continuing education, parental participation, and student access to educational materials. Educational facilities and infrastructure (seven items) covered the condition of the school building,

the suitability of classrooms, access to textbooks, the library, laboratory facilities, electricity, and access to educational technologies. Teacher quality and teaching process (seven items) covered regular teacher attendance, subject expertise, teaching quality, the use of appropriate teaching methods, attention to students, regular evaluation, and feedback. Geographical accessibility (five items) covered the distance from home to school, travel time, route difficulty, access to transportation, and the geographical conditions of the route; items on distance, travel time, and route difficulty were reverse-coded so that higher scores consistently indicate better accessibility. The direction of the accessibility coding must be verified against the original scoring key.. Environmental conditions (five items) covered the safety of the route to school, the safety of the school environment, the stability of educational conditions, the suitability of the social environment, and students' sense of security. Educational status (five items) covered academic progress, motivation to continue education, regular attendance, satisfaction with the educational situation, and achievement of educational goals.

Educational status is therefore operationalized as a composite of students' self-reported perceptions of their schooling situation, not as an objective measure of achievement such as test scores or grades. For each construct, the mean of its items was calculated, so all indicators range from 1 to 5, with higher values indicating a more favorable situation. The predictors are denoted X_1 (economic status and family support), X_2 (educational facilities and infrastructure), X_3 (teacher quality and teaching process), X_4 (geographical accessibility), and X_5 (environmental conditions), and the outcome is denoted Y .

2.3 Statistical Models

The analysis proceeded in four steps. First, descriptive statistics, Cronbach's alpha coefficients, Pearson correlations, and variance inflation factors (VIF) were computed. Second, a classical linear regression was fitted as a baseline:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \varepsilon_i \quad (1)$$

where ε_i is the residual for student i . Third, a single-level Bayesian linear regression (Model 1) was fitted with the likelihood and linear predictor

$$Y_i \sim N(\mu_i, \sigma^2), \quad \mu_i = \beta_0 + \sum_{k=1}^5 \beta_k X_{ki} \quad (2)$$

Fourth, to account for the clustering of students within schools, a Bayesian hierarchical model with school-specific random intercepts and district indicators (Model 2) was fitted. For student i in school j ,

$$Y_{ij} \sim N(\mu_{ij}, \sigma^2), \quad \mu_{ij} = \alpha_j + \sum_{k=1}^5 \beta_k X_{kij} \quad (3)$$

$$\alpha_j = \beta_0 + \gamma_1 D_{1j} + \gamma_2 D_{2j} + u_j, \quad u_j \sim N(0, \sigma_u^2) \quad (4)$$

where D_{1j} and D_{2j} indicate schools in Dawlatyar and Lal wa Sarjangal, respectively, and u_j are school random effects assumed to be independent of the student-level residuals. Firozkoh was chosen as the reference district because it is the provincial capital and contributes the largest subsample ($n = 172$). The predictors X_1 to X_5 are student-level variables (each student's ratings), whereas the district indicators are school-level variables. The centering specification and any school-mean covariates must be confirmed from the analysis script.

The following priors were specified for both models:

$$\beta_k \sim N(0, 10^2), k = 0, \dots, 5; \quad \gamma_1, \gamma_2 \sim N(0, 10^2) \quad (5)$$

$$\sigma \sim \text{Half-Normal}(0, 5); \quad \sigma_\alpha \sim \text{Half-Normal}(0, 1) \quad (6)$$

The normal priors in Equation 5 apply to the intercept, the five regression slopes, and the two district coefficients, each with mean 0 and standard deviation 10. Equation 6 assigns a half-normal prior with scale 5 to the residual standard deviation and a half-normal prior with scale 1 to the between-school standard deviation. These priors have broad support relative to the 1–5 measurement scale. In particular, the residual prior permits values well beyond the observed outcome standard deviation of 0.74, so prior predictive assessment is needed before characterizing the priors as appropriately informative. The draft reports sensitivity analyses using normal slope priors with standard deviations 1 and 100, with unchanged coefficient directions; quantitative comparisons of posterior means, credible intervals, and variance components are needed to substantiate robustness. The posterior distribution is proportional to the product of the likelihood and priors:

$$p(\boldsymbol{\theta} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \quad (7)$$

where $\boldsymbol{\theta}$ denotes all model parameters and $\mathbf{y}$ the observed educational status scores. A Gaussian likelihood was used for the educational status score because it is the mean of five Likert items and is therefore approximately continuous; its observed distribution (range 1.40 to 4.80, mean 3.21, SD 0.74; Table 2) does not reach the limits of the scale, so boundary effects are expected to be limited. Nevertheless, a Gaussian model can, in principle, predict values outside the 1–5 range, and this was examined with posterior predictive checks. No alternative-likelihood results are supplied; sensitivity to the Gaussian specification remains a limitation.

2.4 Computation and Model Assessment

Posterior distributions were estimated using Markov chain Monte Carlo. The draft describes the No-U-Turn Sampler (NUTS), a Hamiltonian Monte Carlo algorithm implemented in Stan (Carpenter et al., 2017), which can be accessed through interfaces such as brms (Bürkner, 2017). The exact software and version, number of chains, warm-up and retained iterations, random seed, and sampler settings must be verified

from the analysis scripts. Convergence was assessed using the rank-normalized R-hat statistic and effective sample sizes (Vehtari et al., 2021).

The models were compared with Pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO) and the widely applicable information criterion (WAIC) (Vehtari et al., 2017). LOO results are reported on the deviance scale,

$$\text{LOOIC} = -2 \widehat{\text{elpd}}_{\text{loo}}, \quad \widehat{\text{elpd}}_{\text{loo}} = \sum_{i=1}^n \log p(y_i | \mathbf{y}_{-i}) \quad (8)$$

so that lower values indicate better expected out-of-sample predictive performance. Finally, graphical posterior predictive checks were used to compare data replicated from the fitted model with the observed data (Gabry et al., 2019).

2.5 Ethical Considerations

The manuscript describes voluntary student participation, confidentiality, and research-only use of the data. Reporting of research permission, the responsible authority, any approval reference and date, and the assent and consent procedures requires confirmation from the original study records. Because some participants may have been minors, the final account must specify how their assent and parental or guardian consent were handled. The authors should report only procedures actually undertaken and clarify how identifying information and access to the data were managed.

2.6 Data Status

The reported study concerns questionnaire data from 372 students collected in 2019. This revision retains the numerical summaries supplied in the manuscript and uses those summaries to complete interpretable narrative passages and arithmetic comparisons. The raw dataset, posterior draws, and analysis scripts were not supplied for this revision; therefore, the reported estimates have not been independently reproduced, and missing empirical results have not been replaced with invented values. The authors should verify all reported results against the original dataset and analysis outputs before submission.

3. Results

3.1 Demographic Characteristics of Respondents

All 372 returned questionnaires were checked for completeness, and none was excluded; the final sample therefore comprised 372 students. Of these, 198 (53.23%) were boys and 174 (46.77%) were girls; 172 students (46.24%) attended schools in Firozkoh, 86 (23.12%) in Dawlatyar, and 114 (30.65%) in Lal wa Sarjangan (Table 1). The mean age was 17.2 years (SD = 1.14), and 189 students were in Grade 11 and 183 in Grade 12.

Table 1

Demographic Characteristics of the Respondents (N = 372)

Variable	Category	Frequency	Percentage
Gender	Male	198	53.23
	Female	174	46.77
Grade	Grade 11	189	50.81
	Grade 12	183	49.19
District	Firozkoh	172	46.24
	Dawlatyar	86	23.12
	Lal wa Sarjangan	114	30.65
Total		372	100.00

3.2 Descriptive Statistics and Reliability

Table 2 reports the descriptive statistics of the research variables. Because all indices are based on a five-point scale, their theoretical range is 1 to 5. The mean educational status score was 3.21 (SD = 0.74). Among the predictors, teacher quality and teaching process had the highest mean (3.18), whereas geographical accessibility (2.59) and educational facilities and infrastructure (2.64) had the lowest means, suggesting that access to school and the condition of school infrastructure are the least favorable aspects of schooling in the studied districts.

Table 2

Descriptive Statistics of the Research Variables (N = 372)

Variable	N	Mean	SD	Minimum	Maximum
Economic status and family support (X_1)	372	2.87	0.81	1.00	5.00
Educational facilities and infrastructure (X_2)	372	2.64	0.92	1.00	4.86
Teacher quality and teaching process (X_3)	372	3.18	0.83	1.14	5.00
Geographical accessibility (X_4)	372	2.59	0.95	1.00	5.00
Environmental conditions (X_5)	372	3.02	0.87	1.00	5.00
Educational status (Y)	372	3.21	0.74	1.40	4.80

Cronbach's alpha was 0.91 for the full questionnaire and ranged from 0.77 to 0.89 across the constructs (Table 3), indicating acceptable to good internal consistency. These coefficients do not establish content or construct validity. In addition, the full-questionnaire alpha should be interpreted cautiously because the questionnaire combines six distinct constructs; the construct-specific coefficients are more relevant to the composite scores used in the analysis.

Table 3

Reliability of the Research Constructs

Construct	Number of items	Cronbach's alpha
Economic status and family support	6	0.81
Educational facilities and infrastructure	7	0.87
Teacher quality and teaching process	7	0.89
Geographical accessibility	5	0.77
Environmental conditions	5	0.83
Educational status	5	0.86
Total questionnaire	35	0.91

3.3 Correlations and Collinearity

Educational status was positively correlated with all five predictors (Table 4). The strongest correlation was with teacher quality and teaching process ($r = 0.62$), followed by educational facilities and infrastructure ($r = 0.56$). All VIF values were below 2 (range 1.29–1.74; Table 5), well below the conventional threshold of 5, indicating no problematic multicollinearity; all five predictors were therefore retained.

Table 4

Correlation Matrix of the Research Variables

Variable	X_1	X_2	X_3	X_4	X_5	Y
X_1 Economic	1.00	0.35	0.28	0.22	0.31	0.41
X_2 Facilities	0.35	1.00	0.46	0.33	0.42	0.56
X_3 Teachers	0.28	0.46	1.00	0.27	0.38	0.62
X_4 Geography	0.22	0.33	0.27	1.00	0.29	0.31
X_5 Environment	0.31	0.42	0.38	0.29	1.00	0.37
Y Educational status	0.41	0.56	0.62	0.31	0.37	1.00

Table 5

Variance Inflation Factors of the Predictors

Predictor	VIF
Economic status and family support	1.38
Educational facilities and infrastructure	1.74
Teacher quality and teaching process	1.62
Geographical accessibility	1.29
Environmental conditions	1.56

3.4 Classical and Single-Level Bayesian Regression

In the classical regression (Equation 1), teacher quality and teaching process had the largest positive coefficient, and educational facilities and economic status and family support were also positively associated with educational status. Because the purpose of the study was probabilistic inference and the quantification of uncertainty, the Bayesian models were used for the main analysis. Table 6 presents the posterior estimates of Model 1. All five coefficients were positive. Teacher quality and teaching process showed the largest posterior mean coefficient (0.31; 95% CrI [0.21, 0.41]), followed by educational facilities and infrastructure (0.26; 95% CrI [0.16, 0.36]), environmental conditions (0.17), economic status and family support (0.14), and geographical accessibility (0.09). The posterior probability of a positive coefficient was at least 0.98 for every predictor.

Table 6

Posterior Estimates of the Single-Level Bayesian Regression Model (Model 1)

Parameter	Posterior mean	Posterior SD	95% CrI	$P(\beta > 0 y)$
β_0 Intercept	1.08	0.21	[0.68, 1.49]	0.99
β_1 Economic status and family support	0.14	0.05	[0.04, 0.24]	0.99
β_2 Educational facilities	0.26	0.05	[0.16, 0.36]	1.00
β_3 Teacher quality	0.31	0.05	[0.21, 0.41]	1.00
β_4 Geographical accessibility	0.09	0.04	[0.01, 0.17]	0.98
β_5 Environmental conditions	0.17	0.05	[0.07, 0.27]	0.99

Note. CrI = credible interval. Coefficients are unstandardized. Values of 1.00 in the last column are rounded posterior probabilities (> 0.995) and do not indicate certainty.

All predictors are described as composite mean scores on a 1–5 metric, and Table 6 reports unstandardized coefficients. Each slope therefore represents the expected difference in perceived educational status associated with a one-point difference in the predictor, holding the other predictors constant. Any centering affects the interpretation of the intercept and must be confirmed from the analysis script. Teacher quality has the largest posterior mean, but its 95% credible interval overlaps that of educational facilities. A probability that one coefficient exceeds the other requires joint posterior draws and cannot be recovered exactly from the marginal summaries in Table 6.

3.5 Convergence Diagnostics

The manuscript reports R-hat values of 1.00 after rounding for all parameters and effective sample sizes exceeding 2,000 for the main parameters. These summaries suggest adequate mixing for the reported parameters, but a complete assessment requires unrounded maximum R-hat and minimum bulk and tail effective sample sizes across both models, including variance components. Counts of divergent transitions

and maximum-tree-depth hits, energy diagnostics, and trace and rank plots are not supplied and should be checked in the original sampler output.

3.6 Hierarchical Model, Model Comparison, and Posterior Predictive Checks

In Model 2, the posterior mean between-school standard deviation was 0.28, indicating residual school-level heterogeneity after adjustment for the predictors and district. Table 7 shows that LOOIC decreased from 786.1 to 749.3, an improvement of 36.8 points, equivalent to an increase of 18.4 in expected log predictive density. WAIC decreased from 785.4 to 748.7, an improvement of 36.7 points. Both criteria favor the hierarchical model. The standard error of the paired predictive difference and Pareto k diagnostics are not reported, so the uncertainty and reliability of this comparison cannot be fully assessed. Student-level PSIS-LOO evaluates prediction for new students in the existing schools; prediction for new schools requires leave-one-school-out validation. The hierarchical model was retained because it represents the clustered design and has lower reported predictive criteria.

Table 7

Comparison of the Bayesian Models Based on LOOIC and WAIC

Model	LOOIC	WAIC
Model 1: Single-level Bayesian regression	786.1	785.4
Model 2: Bayesian hierarchical regression	749.3	748.7

Note. Both criteria are on the deviance scale; lower values indicate better expected out-of-sample predictive performance.

Graphical posterior predictive checks are described in the draft as showing broad agreement between observed and replicated educational status distributions, with discrepancies at the extremes. The corresponding plots and numerical summaries are not supplied, so this assessment remains to be verified. Useful checks include observed and replicated score distributions, the proportion of predictions outside the 1–5 scale, and observed versus replicated school means. Extreme-score discrepancies may reflect the Gaussian likelihood, omitted predictors, or measurement limitations; they do not by themselves identify an omitted cause.

3.7 School-Specific Effects and District Differences

The reported posterior means of the 32 school-specific effects range from approximately -0.42 to $+0.47$. These are adjusted deviations from the district-specific regression intercept, rather than unadjusted school means. Together with the reported between-school standard deviation of 0.28, their spread suggests residual school heterogeneity. School-specific credible intervals and the number excluding zero are not available in the manuscript; therefore, individual schools cannot be classified as clearly above or below the conditional baseline. A plot of all 32 effects and their 95% credible intervals requires the original posterior output.

Table 8 reports district coefficients relative to Firozkoh. The posterior mean difference is -0.07 for Dawlatyar (95% CrI $[-0.21, 0.07]$) and -0.12 for Lal wa Sarjangan

(95% CrI [-0.25, 0.01]). Both intervals include zero, leaving uncertainty about the direction of each adjusted district association. This does not establish district equivalence. Exact posterior probabilities of negative district coefficients require posterior draws and cannot be inferred exactly from these intervals. The reported coefficients suggest modest adjusted district differences, while the hierarchical model also allows variation among schools within districts.

Table 8

Posterior Mean Differences Between Districts (Model 2; Reference = Firozkoh)

District	Posterior mean difference	95% CrI
Firozkoh (reference)	—	—
Dawlatyar	-0.07	[-0.21, 0.07]
Lal wa Sarjantal	-0.12	[-0.25, 0.01]

4. Discussion

4.1 Interpretation of the Findings

This study shows that students' educational status in less developed areas of Ghor Province is associated with several factors rather than with a single one. Teacher quality and teaching process had the largest posterior mean coefficient after adjustment for the other predictors, although its credible interval overlapped with that of educational facilities. In settings where access to educational resources and technology is limited, the teacher's role in explaining content, providing feedback, and sustaining attendance may become more prominent. This interpretation is consistent with studies identifying school and teacher factors as major correlates of academic performance (Tilbe & Biru, 2026) and with reviews of educational interventions in developing countries, which repeatedly identify pedagogy and teacher-related programs among the more promising approaches, despite differences in their conclusions (Evans & Popova, 2016). However, because teacher quality was rated by the students themselves, part of this association may reflect a general tendency of students who view their schooling positively to rate both their teachers and their educational situation more favorably.

Educational facilities and infrastructure had the second-largest coefficient, and its 95% credible interval lay entirely in the positive range, indicating that access to textbooks, adequate classrooms, laboratories, electricity, and educational technology is associated with a more favorable educational status. Economic status and family support was also positively associated with educational status, which is consistent with the view that families with greater economic resources are better able to meet students' educational needs and to support continued schooling, and with research linking income, education, and other poverty-related variables (Gómez-Méndez &

Amornbunchornvej, 2024). Environmental conditions were positively associated with educational status as well, suggesting that a safe and stable environment provides a basis for regular attendance and continued learning.

Geographical accessibility had the smallest coefficient, although its association remained positive. This does not imply that accessibility is unimportant; rather, its unique association, conditional on the other predictors, was smaller, possibly because part of its role operates through school facilities and environmental conditions. Methodologically, the hierarchical model showed that part of the variation in educational status lies between schools and that accounting for this structure improves expected predictive performance, in line with previous applications of Bayesian hierarchical models to educational data (Mosia et al., 2025; Tomal et al., 2021). The posterior distributions also allowed uncertainty to be communicated directly through credible intervals and posterior probabilities rather than through point estimates alone (Kruschke & Liddell, 2018).

4.2 Limitations and Future Research

These findings should be interpreted with caution. First, the data are cross-sectional and were obtained from a convenience sample in three districts, so the estimated coefficients represent conditional associations rather than causal effects, and their generalizability to all students in Ghor Province or to other regions is limited. Second, unmeasured confounders, such as household composition, students' prior achievement, or local security conditions, may be related to both the predictors and educational status. Third, educational status was measured as a composite of students' perceptions rather than objective achievement, and all variables were self-reported at a single time point, which may inflate correlations through common-method and perception bias. Fourth, a Gaussian likelihood was used for a bounded 1–5 composite score, the model comparison evaluated prediction for new students in existing schools only, and the full posterior summary of the hierarchical model (including the residual variance and the intraclass correlation) and quantitative prior-sensitivity results are not reported, which limits the assessment of the magnitude of between-school variation and of robustness to the priors. Fifth, the data were collected in 2019, before the major political and educational changes in Afghanistan since August 2021, including restrictions on girls' secondary education; the results therefore describe pre-2021 conditions and may not reflect the current situation in Ghor Province.

Future studies should use probability sampling, longitudinal data, and objective achievement measures alongside perception-based indicators and, where the number of regions allows, a third hierarchical level for regions. Multilevel Bayesian models with more complex structures could also separate student, family, school, and regional effects more fully.

5. Conclusion

This study modeled the factors associated with the educational status of students in less developed areas of Ghor Province, using data from 372 students in 32 schools in Firozkoh, Dawlatyar, and Lal wa Sarjangal and a Bayesian regression and hierarchical modeling framework. Teacher quality, educational facilities, environmental conditions, economic status and family support, and geographical accessibility were all positively associated with educational status, with teacher quality showing the largest posterior mean coefficient.

Methodologically, the Bayesian hierarchical model showed better expected predictive performance than the single-level model, and accounting for school-specific effects allowed between-school heterogeneity to be taken into account. The posterior distributions made it possible to examine the uncertainty of each parameter and the probability of each association being positive, rather than relying solely on point estimates, which makes the Bayesian approach well suited to educational analysis in settings characterized by uncertainty and heterogeneity.

Practically, the findings suggest that improving teacher quality and strengthening school facilities deserve priority in educational planning for less developed regions, alongside providing a safe environment for students, supporting families facing economic hardship, and reducing geographical barriers to school access. Because the findings are correlational, the effectiveness of such measures should be tested in longitudinal or intervention studies.

Statements and Declarations

Funding:

Funding information was not provided in the manuscript and must be confirmed by the authors before submission.

Competing Interests:

The draft states that the authors have no competing interests; all authors should confirm this declaration before submission.

Ethics Statement:

The final ethics statement must identify the permission or approval actually obtained and accurately describe student assent, consent for minors, and confidentiality procedures. These details are not documented in the supplied manuscript.

Data Availability:

A repository link or a confirmed procedure for accessing the anonymized dataset and analysis code has not been provided. The authors should specify whether and under what conditions the materials can be shared, taking account of participant consent and confidentiality.

Author Contributions:

Individual CRediT roles require confirmation by the named authors and cannot be assigned from the manuscript alone.

Declaration of Generative AI Use:

Generative AI assisted with language editing, editorial restructuring, and the completion of narrative passages using numerical summaries already present in the manuscript. It did not generate missing field data or posterior estimates. The authors are responsible for verifying the revised text against the original data and analysis outputs and approving the final manuscript.

Acknowledgements:

The authors thank the participating students and schools in Firozkoh, Dawlatyar, and Lal wa Sarjantal for their participation in the study.

6. References

- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32. <https://doi.org/10.18637/jss.v076.i01>
- Depaoli, S., & van de Schoot, R. (2017). Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological Methods*, 22(2), 240–261. <https://doi.org/10.1037/met0000065>
- Evans, D. K., & Popova, A. (2016). What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. *The World Bank Research Observer*, 31(2), 242–270. <https://doi.org/10.1093/wbro/lkw004>
- Fouskakis, D., Petrakos, G., & Vavouras, I. (2014). *A Bayesian hierarchical model for comparative evaluation of teaching quality indicators in higher education* (arXiv:1404.1710). arXiv. <https://arxiv.org/abs/1404.1710>
- Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182(2), 389–402. <https://doi.org/10.1111/rssa.12378>
- Gómez-Méndez, I., & Amornbunchornvej, C. (2024). Income, education, and other poverty-related variables: A journey through Bayesian hierarchical models. *Heliyon*, 10(6).
- Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, 25(1), 178–206. <https://doi.org/10.3758/s13423-016-1221-4>
- Le, F., Stanford, T. E., Dumuid, D., & Wiley, J. F. (2025). Bayesian multilevel compositional data analysis: Introduction, evaluation, and application. *Psychological Methods*.

- McNeish, D. (2016). On using Bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(5), 750–773. <https://doi.org/10.1080/10705511.2016.1186549>
- Montano, V., & Reyes, A. (2026). Beyond course averages: A generalized Bayesian hierarchical framework for course-level learning evaluation. *Journal of Learning Theory and Methodology*, 7(1), 37–48.
- Mosia, M., Egara, F. O., Nannim, F. A., & Basitere, M. (2025). Bayesian hierarchical modelling of student academic performance: The impact of mathematics competency, institutional context, and temporal variability. *Education Sciences*, 15(2), Article 177. <https://doi.org/10.3390/educsci15020177>
- Mosia, M., Sekonyela, L., Egara, F. O., Nimy, E., Mabokgole, I. M., & Nannim, F. A. (2026). Bayesian hierarchical modelling of academic orientation and advising effects on student retention and progression: Multi-cohort evidence. *PLOS One*, 21(3), Article e0345001. <https://doi.org/10.1371/journal.pone.0345001>
- Pardo, L., Sosa, J., Torres-Clavijo, J. P., & Arévalo-Arévalo, A. F. (2025). *Uncovering latent territorial structure in ICFES Saber 11 performance with Bayesian multilevel spatial models* (arXiv:2512.17119). arXiv. <https://arxiv.org/abs/2512.17119>
- Sun, Y. (2026). Using hierarchical Bayesian analysis to identify at-risk students in engineering mathematics prerequisites. *Computer Applications in Engineering Education*, 34(1), Article e70126. <https://doi.org/10.1002/cae.70126>
- Tilbe, Y. T., & Biru, D. (2026). Determinants of academic performance among public and private middle-level school students in Yabelo Town, Ethiopia: Examining the influence of school, teacher, and leadership factors. *Ethiopian Journal of Business and Social Science*, 9(1), 18–46.
- Tomal, J., Rahmati, S., Boroushaki, S., Jin, L., & Ahmed, E. (2021). The impact of COVID-19 on students' marks: A Bayesian hierarchical modeling approach. *METRON*, 79(1), 57–91. <https://doi.org/10.1007/s40300-021-00200-1>
- van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märten, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1, Article 1. <https://doi.org/10.1038/s43586-020-00001-2>
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis*, 16(2), 667–718. <https://doi.org/10.1214/20-BA1221>
- Wafa, M. N., Hussaini, S. A. M., & Pazhman, J. (2020). Evaluation of students' mathematical ability in Afghanistan's schools using cognitive diagnosis models.



Eurasia Journal of Mathematics, Science and Technology Education, 16(6), Article em1849. <https://doi.org/10.29333/ejmste/7834>

Wafa, M. N., & Shuja, S. (2024). Optimizing variable selection with Bayesian t-Lasso regression: A comparison of normal-mixture and uniform-mixture representations for outlier data analysis. *Polyhedron International Journal in Mathematics Education*, 2(1), 17–29.

Wafa, M. N., Zia, Z., & Hussaini, S. A. M. (2023). Regression models according to Birnbaum–Saunders distribution. *European Journal of Mathematics and Statistics*, 4(6), 24–30.



Author verification before submission

Author names, affiliations, corresponding email, and article-history dates must be entered from verified records. They cannot be inferred from the article.

Confirm school selection and recruitment, district-specific school counts, school sample sizes, instrument language and adaptation, content review, pilot testing, and construct validity.

Confirm that the instrument was developed before the 2019 survey. A 2026 publication cannot substantiate the original instrument-development process; it may only be used for retrospective discussion.

Check the regression intercept, predictor centering, and Table 6 against the original output. Using the reported predictor means and slopes gives $1.08 + 0.14(2.87) + 0.26(2.64) + 0.31(3.18) + 0.09(2.59) + 0.17(3.02) = 3.9009$, whereas the reported outcome mean is 3.21. Under the uncentered specification, an intercept of approximately 0.3891 would align these summaries. This is a consistency check, not a replacement estimate; confirm the model specification and values before changing the table.

Confirm the actual priors, centering, software version, sampler settings, sensitivity analyses, unrounded diagnostics, and predictive comparisons. Missing standard errors and posterior probabilities require analysis outputs, not arithmetic reconstruction.

Provide full Model 2 coefficient summaries, residual and school variance intervals, and posterior intraclass correlation. The ICC is $\sigma_{\text{school}}^2 / (\sigma_{\text{school}}^2 + \sigma_{\text{residual}}^2)$; it cannot be computed from the school standard deviation alone.

Provide posterior predictive plots and school-effect intervals if available. Do not create empirical plots from approximate or invented posterior draws.

Confirm funding, competing interests, actual ethical permission and consent procedures, data-sharing arrangements, author contributions, and the acknowledgement wording.