*American Journal of Educational Research*, 2016, Vol. 4, No. 14, 1030-1040
Available online at http://pubs.sciepub.com/education/4/14/8

DOI:10.12691/education-4-14-8

# A Hierarchical Linear Modelling of Teacher Effects on Academic Achievement in the Kenya Certificate of Primary Education Examination

**Epari Ejakait[1,*], Maureen Olel[2], Lucas Othuon[3], Ochanda Khasenye[4]**

[1]Department of Educational Planning and Management, Masinde Muliro University of Science and Technology, Kakamega, Kenya
[2]Department of Educational Management and Foundations, Maseno University, Private Bag, Maseno, Kenya
[3]Department of Educational Psychology, Maseno University, Private Bag, Maseno, Kenya
[4]Department of Environment and Health Sciences, Technical University of Mombasa, P.O Box 90420-80100, GPO, Mombasa, Kenya
*Corresponding author: eejakait@mmust.ac.ke, eejakait@gmail.com

**Abstract** The past five decades have seen rapid expansion in academic achievement surveys with mixed findings and interpretation. Utilizing the education production function models, the surveys sought to test whether school or teacher-level variables explain academic achievement variance to a greater extent than student-level variables. Within this framework, we modelled teacher-level predictors of academic achievement in the Kenya Certificate of Primary Education (KCPE) examination in Mumias and Kuria East Sub-Counties in Kenya. Using a three-level hierarchical linear model (with 1824 students at Level-1 nested within 305 teachers at Level-2 who were themselves nested within 61 schools at Level-3), the results suggest that adjusting for Level-1 and Level-3 covariates, teacher age, the number of short in-service courses attended by the teachers in their respective subject areas and the number of formal written tests in those respective academic subjects have statistically significant effect on student academic achievement in the Kenya Certificate of Primary Education Examination. Policy implications of these findings are discussed.

***Keywords:*** *teacher-level predictors, hierarchical linear modelling, Kenya Certificate of Primary Education Examination, Mumias Sub-County, Kuria East Sub-County, Kenya*

**Cite This Article:** Epari Ejakait, Maureen Olel, Lucas Othuon, and Ochanda Khasenye, “A Hierarchical Linear Modelling of Teacher Effects on Academic Achievement in the Kenya Certificate of Primary Education Examination.” *American Journal of Educational Research*, vol. 4, no. 14 (2016): 1030-1040. doi: 10.12691/education-4-14-8.

## 1. Introduction

The large body of literature on academic achievement since the Coleman Report on the Equality of Educational Opportunity in the US [1] probably shows the desire by scholars to understand and clarify the drivers and predictors of student academic achievement. The Coleman Report made some landmark findings that spurred lots of research and interest in education production functions and schooling outcomes. An unexpected finding from the study that took many policy analysts by surprise suggested that variations in the level of students' achievements bore little or no relationship with school resources or programmes. That instead, out-of-school variables such as family background and neighbourhood characteristics accounted for the observed achievement differences between students.

Utilizing the education production function models, many studies were designed following the findings of the Coleman Report. The studies sought to measure whether school or teacher-level variables were able to explain academic achievement variance to a greater extent than student-level characteristics. Three strands of findings seem to have emerged from this large body of studies since 1966: That school-level variables explain a large proportion of the variance in student academic achievement [2,3,4,5,6]; that teacher-level variables have a prominent effect on students achievement [7,8] and that student-level variables including their socioeconomic status and other home background characteristics account for much of the variation in their academic achievement [1,9].

This paper pursues the teacher-level effect using a three-level hierarchical linear model with 1824 students (Level-1) nested within 305 teachers (Level-2) who are themselves nested within 61 primary schools (Level-3). We test the hypothesis that teacher-level variables have no effect on student academic achievement in the five academic subjects offered in the Kenya Certificate of Primary Education (KCPE) examination. These are English, Kiswahili, Mathematics, Science and Social Studies and Religious Education with each scored out of a possible 100. The KCPE examination is taken at the end of eight years of primary education which marks the end of the primary education cycle and opens way for transition into secondary school education. We now review selected literature on teacher-effects.

While examining the effect of primary school quality on academic achievement across 29 high and low income countries in Africa, Asia, Latin America and the Middle

East, Heyneman and Loxley [10,11] concluded that teacher and school quality variables were the most important in influencing student learning and academic achievement. They also argued that ''...the poorer the national setting in economic terms, the more powerful this school and teacher quality effect appears to be...'' [[11], p. 1184]. Their conclusion, commonly referred to as the ''Heyneman–Loxley effect'' or ''HL effect'' in educational literature was considered important because it supported the linkage between educational achievement and national economic development [5].

The work of Sanders and Rivers [12] as well as Wright *et.al.* [13] using a database of approximately 3 million records for Tennessee's entire grade 2-6 student population for the period between 1990 and 1996 involving student achievement scores in mathematics, reading, language arts, science, and social studies, indicated that teacher effects were highly significant in 20 of the 30 analyses done and had larger effect sizes than any other factor. Using a 50-state survey of policies in the US in 94 schools from Staffing Surveys and the National Assessment of Educational Progress, Darling-Hammond's findings [14] suggest that teacher preparation and certification were the strongest correlates of student achievement in reading and mathematics, both before and after controlling for student poverty and language status.

Still on teacher-effects, Laczko-Kerr and Berliner [15] found that the students of certified teachers outperformed students of under-certified teachers by about 2 months on a grade equivalent scale in reading, mathematics, and language. Further, the students of under-certified teachers made about 20% less academic growth per year compared with that made by students of teachers with regular certification. Wenglinsky [16] applied Multilevel Structural Equation Modelling on the 1996 National Assessment of Educational Progress in mathematics in the US and concluded that the effects of classroom practices were comparable in size to those of student background especially when added to those of other teacher characteristics. Using HLM on data from a four-year experiment involving teachers and students, Nye, Hedges and Konstantopoulos [17] found that teachers had larger effects on achievement in mathematics than in reading and that there was a substantial relationship between teacher experience and student achievement gains. Rivkin et al. [18] also found that teacher variables had huge effects on achievement in reading and mathematics. Using structural equation modelling, Caprara, Barbaranelli, Steca, and Malone [19] examined teacher's self-efficacy beliefs and job satisfaction as well as student's academic achievement aggregated at the school level, focusing on indicators of school functioning. A sample of 2184 teachers in 75 Italian junior high schools was administered self-report questionnaires to assess self-efficacy beliefs and their job satisfaction. Standardized final examination grades at the end of the third year of junior high school were used to assess the students' achievement. The findings suggested that previous student's academic achievement predicted subsequent achievement as well as teacher's self efficacy beliefs, which, in turn, contributed significantly to student's achievement and teacher's job satisfaction [19].

Neild [7] employed a three-level HLM on data created from a student report card and administrative files merged with human resource files on teacher characteristics from an urban district in the USA to estimate the impact of different teacher certification categories on middle-grades students' learning gains in mathematics and science. The student sample comprised of grades 5 through 8 in all public, non-charter middle schools in the 2002-2003 school year. The findings suggested that in mathematics, students with elementary and secondary-certified teachers outscored those with uncertified teachers and those whose teachers were certified in special education. Strong effects were seen in science, where students with secondary science-certified teachers substantially outscored those with any other kind of teacher [7].

An analysis of data from Project STAR and the Lasting Benefits Study in the USA by Konstantopoulos and Chung [8] using a three-level hierarchical linear model examined whether teacher effects from kindergarten to fifth grade could simultaneously affect sixth grade achievement. Their findings demonstrated that teachers affect student achievement positively and that teacher effects persist through the sixth grade for mathematics, reading, and science achievement. Further, they suggest that this cumulative effect of teachers on student achievement was considerable.

Although studies as reviewed in this sub-section have found significant teacher effects on student academic achievement, some renowned scholars in Economics of Education such as Eric Hanushek, have reported a lack of significant positive relationship between school-level variables and student achievement. For instance, Hanushek [20] found that estimated coefficients for teacher-based variables were statistically insignificant and that there was no strong evidence suggesting that Pupil Teacher Ratio (PTR), teacher education or teacher experience had an expected positive effect on student achievement. Similar results are reflected from Hanushek's review of 187 studies on expenditure relationships in schools [21]. From the studies, only 14 out of 152 that dealt with effects of class size reported statistically significant relationships. On teacher education, 100 out of 113 studies showed statistically insignificant coefficients. The conclusion from this review argued that there was no strong evidence that PTR, teacher education, or teacher experience had the expected positive effects on student achievement and that there was no strong or systematic relationship between school expenditures and student performance [21]. In another review of close to 400 studies on student achievement, Hanushek [22] argued that after accounting for family inputs, this large body of literature did not present compelling or consistent results that suggest a relationship between student performance and school resources.

So which results should readers believe? The answer probably lies in data handling and the choice of statistics used to generate the findings one is interested in. We argue that it is prudent to choose HLM over Ordinary Least Squares (OLS) for hierarchical data such as is commonly found in the modelling of academic achievement. Available literature probably presents three limitations that OLS runs into when confronted with hierarchical data because it is not sensitive to such data. Hierarchical Linear Modelling involves relating variables at one level of analysis e.g. the students, to another level of analysis e.g. the teachers or classrooms and to yet another level of analysis e.g. the school [23]. Studies using OLS tend to either aggregate student data to the school level or to disaggregate school data to the student level. The first approach can introduce aggregation biases into the models, the second approach can seriously underestimate standard errors, and both approaches can miss important information about the nature of the school effects [16,23,24].

A second limitation with OLS techniques is their failure to take measurement error into account. These techniques assume that the variables in the models are perfectly measured by the observed data. More often, the operationalization of most variables is subject to substantial error, both because the operationalization does not correspond perfectly to the model (e.g. parents' income as a proxy for socioeconomic status) and because data collection procedures are error-prone [25]. Failing to take measurement error into account can lead to biased estimates of model coefficients [16,26].

A third limitation with OLS techniques is that they are not adept at measuring interrelationships among independent variables. School effects often involve a multi-step process, in which one school characteristic influences another that may, in turn, influence the outcome of interest. While it is possible to run a series of models that regress each independent variable on the others, such models tend to be cumbersome and lack precision in measuring the overall fit of the series of models [25]. Because of these difficulties, school effects research often neglects the indirect effects of various school characteristics [16,27]. Therefore, data handling and choice of statistics probably give us the first-line decision hints we need in a credibility assessment of research findings.

The rest of this paper is organised as follows. Section 2 describes the data and methods while Section 3 presents the results and discussion. Section 4 contains our conclusion and implications for policy while section 5 is our statement of competing interests.

## 2. Data and Methods

### 2.1. Sampling

The data were collected from Mumias and Kuria East Sub-Counties in Kenya. There were 280 Sub-Counties and Municipalities as extracted by the authors from the official KCPE examination results dataset for 2012. Mumias and Kuria East were randomly sampled from Sub-Counties that had consistently been in the top and bottom 5% respectively using merit lists for 2010-2012. Stratified by Public and Private primary schools, we employed Probability Proportion to Size to sample 1824 students (Level-1) nested within 305 teachers (Level-2) who were themselves nested within 61schools (Level-3). While all Class 8 candidates in single streamed schools were included in the sample, one stream at Class 8 in multi-streamed schools was randomly sampled and all its candidates included in the sample as well. One teacher for each of the five KCPE examinable academic subjects in each of the 61 schools was sampled bringing the total to 305. Public and private primary schools are established and maintained out of public and private funds respectively.

### 2.2. The Data and Description of Variables

Instead of running separate models for each of the five academic subjects using two-level hierarchical linear models, the authors built a single long-format data file with 9120 records since each of the 1824 students had one score-record for each of the five academic subjects. A description of the academic subjects is summarised in Table 1.

**Table 1. Description of the KCPE Examination Disaggregated by Subject**

| Name of Exam | | Duration (Minutes) | Number of Items | Multiple Choice? | Maximum Score |
|---|---|---|---|---|---|
| English Section A | Language | 100 | 50 | Yes | 50 |
| English Section B | Composition | 40 | 1 | No | 40 |
| Kiswahili Section A | Language | 100 | 50 | Yes | 50 |
| Kiswahili Section B | Composition | 40 | 1 | No | 40 |
| Mathematics | | 120 | 50 | Yes | 50 |
| Science | | 120 | 50 | Yes | 50 |
| Social Studies and Religious Education | | 135 | 90 | Yes | 90 |

*Note.* English's section A and B are combined for the total score in English. The same happens for Kiswahili.

Kiswahili is the first language of the Swahili people (one of the 300-600 ethnic groups in Africa who speak Bantu languages [28]) and is a lingua franca of the African Great Lakes region and other parts of Southeast Africa, including Tanzania, Kenya, Uganda, Rwanda, Burundi, Mozambique, and the Democratic Republic of the Congo [29]. Kiswahili serves as a national language of four nations: Tanzania, Kenya, Uganda, and the Democratic Republic of the Congo and is also one of the working languages of the African Union and one of the official languages of the East African Community [30].

For Mathematics and Science, each of the 50 items has 2 points with 100 as the maximum score. The total score is 90 points for each of English and Kiswahili and the student's final score is calculated as given in equation (1).

$$\frac{x}{me/mk} * 100 \tag{1}$$

Where $x$ is the student's cumulative score in English's or Kiswahili's sections A (scored out of 50) and B (scored out of 40) and $te$ is the maximum score in English while $tk$ is the maximum score in Kiswahili.

The same applies for Social Studies and Religious Education with the final student's score calculated as given in equation (2)

$$\frac{x}{mssre} * 100 \tag{2}$$

Where $x$ is the student's cumulative score from the 90 items in Social Studies and Religious Education and $mssre$ is the maximum score in the same examination.

School, teacher and student questionnaires were fielded for data collection. Table 2 presents a description of the variables used in the three-level hierarchical linear modelling of teacher effects.

**Table 2. Description of Variables Used in the Analysis of the Data**

| Variable | Variable label | Variable scale | Variable values |
|---|---|---|---|
| s17z | Student's running z-score on the 5 subjects | Interval | -3.07-2.93 |
| s21 | Female student | Nominal | 0=Male; 1=Female |
| s22a | Student's age in years | Interval | 12.18 - 22.18 |
| s23a | Student's years in current school | Interval | 0.56 - 10.70 |
| s27 | Number of times student spoke English in the last 7 days | Ratio | 0 - 7 |
| s313x | Student's Wealth Index (3 Tertiles) | Categorical | 1=High tertile; 2=Middle tertile; 3=Low tertile |
| s332 | Mother has some primary education | Dummy | 0=Otherwise; 1=Mother has some primary education |
| s333 | Mother has completed primary education | Dummy | 0=Otherwise; 1=Mother has completed primary education |
| s334 | Mother has some secondary education | Dummy | 0=Otherwise; 1=Mother has some secondary education |
| s353 | Father has completed primary education | Dummy | 0=Otherwise; 1=Father has completed primary education |
| s356 | Father has completed post secondary training | Dummy | 0=Otherwise; 1=Father has completed post secondary education |
| s36c | Number of siblings | Ratio | 0-16 |
| s53 | How often learners are hurt by teachers | Categorical | 1=Never; 2=Sometimes; 3=Often |
| s58x | Number of times student has repeated classes | Ratio | 0-3 |
| s61x | Student keeps negative company (z-score) | Interval | -0.74 - 2.65 |
| t22a | Teacher's age in years | Interval | 18.21 - 58.07 |
| t214 | Number of in-service courses attended (2013) | Ratio | 0-8 |
| t227 | Number of formal written tests in teacher's subject | Interval | 1-40 |
| h16 | Kuria East Sub-County | Dummy | 0=Mumias; 1=Kuria East |
| h24a | Boarding status at class 8 | Categorical | 1=Day; 2=Boarding; 3=Day and boarding |
| h5122 | School disallows borrowing of library books to take home | Dummy | 0=Otherwise; 1=School disallows borrowing of library books to take home |

*Note.* Student Level-1 variables are prefixed with letter "s", Teacher Level-2 with letter "t" and School Level-3 with letter "s".

For ease of interpretation, the outcome variable was transformed to a standard normal score with a Mean of zero (0) and Standard Deviation of one (1) so that the residuals at each level better approximate the normality assumptions of the models. This transformation allowed the effects of the covariates in the three-level HLM to be interpreted in terms of standard deviation units of our outcome variable [25,31]. The untransformed variable ranged between 4 and 99 with mean score of 52.64 and standard deviation of 15.83.

## 2.3. Model Specification

As is usual for HLM, the starting point was to fit an unconditional model (also called intercept-only, null or empty model) in order to obtain the amounts of variance available for explanation at each level of the hierarchy [5,25]. Consequently, a three-level variance components model was specified and fitted including only an intercept, school and teacher effects, and a student level residual error term. The model did not make any adjustments for predictor variables, only decomposing the total variance in the outcome variable (students' running score on the five KCPE subjects) into separate school, teacher and student variance components. We followed Leckie [31] in specifying the unconditional/null model as:

$$Y_{ijk} = \beta_0 + v_k + u_{jk} + e_{ijk} \tag{3}$$

Assuming that; $v_k \sim N(0, \sigma_v^2)$
$u_{jk} \sim N(0, \sigma_v^2)$
$e_{ijk} \sim N(0, \sigma_v^2)$

Where:

$Y_{ijk}$ is the KCPE academic subject score for student $i$ $(i = 1, \ldots, 9120)$ nested within teacher $j$ $(j = 1, \ldots, 305)$ in school $k$, $(k = 1, \ldots, 61)$;

$\beta_0$ is the mean score across all schools;

$v_k$ is the effect of school $k$;

$u_{jk}$ is the effect of teacher $j$; and

$e_{ijk}$ is the student level residual error term.

The school, teacher effects and the student level residual errors are assumed independent and normally distributed with zero means and constant variances.

Table 3 presents the results of this null model. The random intercept, $\beta_0$, predicts that a student's z-score in any of the Five KCPE examination Subjects will be -0.02 (*SE*=0.09, *p*=.834). Since the outcome variable is approximately normalised, an estimated random intercept of zero, an estimated total variance of approximately one and a non significant intercept are all expected. The random part of the model presents the Variance Partition Coefficient (VPC) for each HLM level. Substituting the Variance Components into equation (4, 5 and 6), the VPC available for explanation at Student ($\sigma_e^2$), Teacher ($\sigma_u^2$) and School ($\sigma_v^2$) levels is 0.4388 (43.88%), 0.0493 (4.93%) and 0.5119 (51.19%) respectively.

$$\sigma_e^2 / (\sigma_e^2 + \sigma_u^2 + \sigma_v^2) \tag{4}$$

$$\sigma_u^2 / (\sigma_e^2 + \sigma_u^2 + \sigma_v^2); \tag{5}$$

$$\sigma_v^2 / (\sigma_e^2 + \sigma_u^2 + \sigma_v^2) \tag{6}$$

The largest variance lay between schools (51.19%) while a substantial one lay among students within teachers (43.88%). Only 4.93% of the variance lay between teachers within schools suggesting that there was only modest variation in the five subjects between teachers. Most of the variation in students' scores was seen between their schools and among themselves.

In adding predictors from the three levels to the unconditional model in equation (3), the authors followed Leckie [31] in specifying the full three-level random intercept slopes model as:

$$
\begin{aligned}
Y_{ijk} &= \beta_0 + \beta_1 s21_{ijk} + \beta_2 s22a_{ijk} + \beta_3 s23a_{ijk} \\
&+ \beta_4 s27_{ijk} + \beta_5 s313x_{ijk} + \beta_6 s332_{ijk} + \beta_7 s333_{ijk} \\
&+ \beta_8 s334_{ijk} + \beta_9 s353_{ijk} + \beta_{10} s356_{ijk} + \beta_{11} s36c_{ijk} \\
&+ \beta_{12} s53_{ijk} + \beta_{13} s58_{ijk} + \beta_{14} s61x_{ijk} + \beta_{15} t22a_{jk} \\
&+ \beta_{15} t214_{jk} + \beta_{15} t227_{jk} + \beta_{16} h16_k + \beta_{17} h24a_k \\
&+ \beta_{17} h5122_k + v_{0k} + v_{1k} s21_{ijk} + v_{14k} s61x_{ijk} + u_{jk} + e_{ijk}
\end{aligned}
\qquad
\begin{aligned}
&\begin{pmatrix} v_{0k} \\ v_{1k} \\ v_{14k} \end{pmatrix} \sim N\left(0, \Omega_v\right), \\
&\Omega_v = \begin{pmatrix} \sigma_{v0}^2 & & \\ \sigma_{v01} & \sigma_{v1}^2 & \\ \sigma_{v014} & \sigma_{v114} & \sigma_{v14}^2 \end{pmatrix}, e_{ijk} \sim N(0, \sigma_e^2)
\end{aligned}
\tag{7}
$$

**Table 3. Three Level Unconditional Model**

| Fixed Effect Variable | Variable label | Null Model | |
|---|---|---|---|
| | | Est. (Std. Err.) | p-value |
| | Intercept, $\beta_{0jk}$ | -0.02 (0.09) | 0.834 |
| Random Effect | | Variance Component | |
| Student (Level-1), $e_{ijk}$ | | 0.4358 (0.01) | |
| Teacher (Level-2), $u_{jk}$ | | 0.0490 (0.01) | |
| School (Level-3), $v_k$ | | 0.5084 (0.09) | |
| Variance Partition Coefficient (VPC) | | | |
| Student (Level-1), $\sigma^2_e$ | | 0.4388 | |
| Teacher (Level-2), $\sigma^2_u$ | | 0.0493 | |
| School (Level-3), $\sigma^2_v$ | | 0.5119 | |
| Model Fit Statistics | | | |
| Deviance | | 18963 | |
| Akaike Information Criterion (AIC) | | 18971 | |
| Bayesian Information Criterion (BIC) | | 19000 | |
| Likelihood Ratio test vs. OLS Regression | | $\chi^2(2) = 6917$ | <.001 |

Note. N= 9120 (1824*5 = 9120, each student has 5 academic subject records); Est. = Estimate; Std. Err. = Standard Error (in parentheses); AIC and BIC statistics = smaller-is-better fit; OLS=Ordinary Least Squares.

A description of these predictors is presented in Table 2.

Two new terms $v_{1k}$ and $v_{114k}$ were added to the model, so that the coefficients of the sex of the student and whether or not student kept negative company became $\beta_{1k} = \beta_1 + v_{1k}$ and $\beta_{14k} = \beta_1 + v_{14k}$ respectively and the community-level variance replaced by a matrix with three new parameters, $\sigma^2_{v0}$ , $\sigma_{v01}$ and $\sigma_{v014}$ . Three random intercept models were fitted in steps starting with Level-1 Student predictors subscripted $ijk$ that were estimated in Model-1. The Level-2 Teacher predictors subscripted $jk$ were added in Model-2 while the Level-3 School predictors with the subscript $k$, were accounted for in Model-3. These predictor variables helped to explain the response variation allocated to the three levels as well as test the hypothesis regarding the relationship between teacher-level predictors and the outcome variable. The slope coefficients of these predictor variables were assumed fixed across Levels 2 and 3.

**Table 4. Non-significant Variables Dropped from the Teacher-Model Only**

| Variable | Variable label | Est. (Std. Err.) | *p*-value |
|---|---|---|---|
| t23 | TSC teacher | -0.07 (0.07) | 0.331 |
| t25 | Teacher's professional grade in 2014 | | |
| 15 | P1 | 0.06 (0.08) | 0.441 |
| 16 | S1/Diploma | 0.08 (0.09) | 0.408 |
| 17 | ATS | 0.14 (0.10) | 0.177 |
| 18 | Graduate | 0.07 (0.09) | 0.445 |
| t213 | Teacher's adequacy of teaching subject 0-10 | -0.00 (0.01) | 0.841 |
| t218 | Teacher lessons per week | -0.01 (0.01) | 0.519 |
| t219 | # of times HT observed teacher teaching | -0.01 (0.01) | 0.143 |
| t220 | # of times DHT observed teacher teaching | 0.01 (0.01) | 0.105 |
| t221 | # of times QUASO has visited teacher | -0.03 (0.02) | 0.093 |
| t224 | School has a Subject resource centre serving it | -0.06 (0.05) | 0.25 |
| t231x | Grade 8 subject specific syllabus coverage | -0.00 (0.00) | 0.862 |
| t232 | Teacher has additional textbooks | -0.04 (0.06) | 0.43 |

*Note.* Std. Err. = Standard Error (in parentheses); TSC=Teachers' Service Commission; P1=Primary 1; S1=Secondary 1; ATS= Approved Teacher Status; HT= Head Teacher; DHT= Deputy Head Teacher; QUASO= Quality Assurance Officer; #=Number.

Models 4 and 5 fitted random slopes because an exploratory analysis indicated that the relationship between the students' running score in the five subjects, the outcome variable (s17z), and student sex ($s21$), 0=Male; 1=Female, and whether or not student kept negative company ($s61x$), standardized score, -0.74 - 2.65, varied across Level-3. In model-4, three teacher-level predictors were omitted, i.e. teacher's age in years (t22a), number of in-service short courses attended by the teacher (t214) and the number of formal written tests in the teacher's subject (t227) but were included in the final Model-5 in order to determine their net value in explained variance.

Selection of “candidate predictors” to be included in the three-level models involved a two-step process informed by the need for parsimony in the final model. In the first step, a pair-wise correlation of all possible variables for each of the three levels was estimated. The second step involved running only those variables that were significantly correlated with the outcome variable in an exploratory Level-specific model while considering the hierarchical nature of the dataset [5,25,32,33]. For the student-level, “candidate predictors” that were correlated with the outcome variable were fitted in a student-only model excluding teacher and school-level predictors. Only statistically significant variables at the 5% level were then preserved as the student-level predictors to be included in subsequent models and levels. This procedure was repeated at teacher and school levels. Table 4 presents the non-significant teacher-level predictors that were dropped leaving only t22a, t214 and t227 for modelling at the school level. STATA version 11.2 was used for data management and analysis with the “xtmixed” command.

# 3. Results and Discussion

## 3.1. Descriptive Statistics of the Variables used in the Modelling

Table 5 presents the descriptive statistics for the variables used in the modeling.

**Table 5. Descriptive Statistics for Variables Used in the Modelling**

| Variable | Variable label | Mean | Standard error (mean) | Standard deviation | Min | Max |
|---|---|---|---|---|---|---|
| s17z | Student's running z-score on the 5 subjects | 0.00 | 0.01 | 1.00 | -3.07 | 2.93 |
| s22a | Student's age in years | 15.27 | 0.01 | 1.31 | 12 | 22 |
| s23a | Student's years in current school | 6.23 | 0.03 | 2.79 | 1 | 11 |
| s27 | Number of times student spoke English in last 7 days | 3.52 | 0.02 | 1.83 | 0 | 7 |
| s36c | Number of siblings | 5.07 | 0.02 | 2.36 | 0 | 16 |
| s61x | Student keeps negative company (z-score) | 0.00 | 0.01 | 0.49 | -0.74 | 2.65 |
| t22a | Class 8 teacher's age in years | 37.88 | 0.10 | 9.40 | 18 | 58 |
| t214 | Number of in-service courses attended (2013) | 0.88 | 0.01 | 1.27 | 0 | 8 |
| t227 | Number of formal written tests in teacher's subject | 11.14 | 0.07 | 6.44 | 1 | 40 |
| s58x | Number of times student has repeated classes | 0.74 | 0.01 | 0.72 | 0 | 3 |
| s21 | Female student | 0=Male | 1=Female | | | |
| | | 927 (50.82) | 897 (49.18) | | | |
| | Dummy Variables | 0=Otherwise | 1=Yes | | | |
| s332 | Mother has some primary education | 1,406 (77.08) | 418 (22.92) | | | |
| s333 | Mother has completed primary education | 1,372 (75.22) | 452 (24.78) | | | |
| s334 | Mother has some secondary education | 1,630 (89.36) | 194 (10.64) | | | |
| s353 | Father has completed primary education | 1,483 (81.30) | 341 (18.70) | | | |
| s356 | Father has completed post secondary training | 1,556 (85.31) | 268 (14.69) | | | |
| h5122 | School disallows borrowing of library books to take home | 1,569 (86.02) | 255 (13.98) | | | |
| s313x | Student's Wealth Index (3 Tertiles) | 1=High tertile | 2=Middle tertile | 3=Low tertile | | |
| | | 608 (33.33) | 608 (33.33) | 608 (33.33) | | |
| s53 | How often teachers hurt learners | 1=Never | 2=Sometimes | 3=Often | | |
| | | 943 (51.70) | 736 (40.35) | 145 (7.95) | | |
| h16 | Kuria East Sub-County | 0=Mumias | 1=Kuria East | | | |
| | | 1,068 (58.55) | 756 (41.45) | | | |
| h24a | Boarding status at class 8 | 1=Day | 2=Boarding | 3=Day and boarding | | |
| | | 1,583 (86.79) | 79 (4.33) | 162 (8.88) | | |

*Note*. *n*=1824; Min=Minimum; Max=Maximum; percentages in parentheses (); Student Level-1 variables are prefixed with letter "s", Teacher Level-2 with letter "t" and School Level-3 with letter "s".

The focus in this paper was to assess the effect of teacher-level variables at Class 8. Three teacher predictors were modelled with the Class 8 teachers' age having a Mean of 37.88 and standard deviation of 9.40, number of in-service courses attended in 2013 (*M*=0.88, *SD*=0.01) and the number of formal written tests in the teacher's subject (*M*=11.14, *SD*=6.44). All interval or ratio predictors under consideration had reasonably small standard errors of the mean suggesting that their calculated means were not quite far away from the true population mean. Using Multiple Correspondence Analysis with non-income or expenditure data as proposed by Filmer and Pritchett [34] and as computed in the Demographic Health Surveys [35,36], the students' wealth index was determined from their reported home ownership of assets, such as cars motor cycles, electronics (including fridges), and bicycles among others; materials used for housing construction; source of lighting; and types of water access and sanitation facilities. This wealth index was then divided into three tertiles of 608 students each categorized as 1=High tertile (wealthiest of the three), 2=Middle tertile and 3= Low tertile (least wealthy of the three).

## 3.2. Bivariate Analysis

Pair-wise correlation was run between the students' running z-score on the Five Subjects (s17z) and the full range of predictors as estimated in the final Model-5. Two of the school-level predictors: Sub-County (h16) and Boarding status at class 8 (h24a) presenting the 'strongest' correlation ($r$= -0.513, $p$<.001) and ($r$= -0.394, $p$<.001) respectively. These were however considered moderate using Taylor's interpretation of correlation coefficients [37]. There was no added value in presenting the entire pair-wise matrix table since the rest of the predictors under consideration including the three teacher variables of interest in the models had weak but statistically significant correlations with the outcome variable: Teachers' age, t22a, ($r$=-0.16, $p$>.001; Number of in-service courses attended, t214, ($r$=0.039, $p$>.001), and Number of formal written tests in the teacher's subject, t227, ($r$=0.349, $p$>.001). No other correlation was stronger than $r$= -0.513.

The results of a an independent t-test with unequal variance, $t$(7885)= 56.62, $p$ <.001, showed that Mumias Sub-County had a z-score of 0.43 standard deviation units above the mean compared with Kuria East Sub-County's -0.61 units below the mean. The strength of the difference between the two z-score means as measured by $R^2$ was 0.29 which is considered a very large effect [38].

A one-way ANOVA was also run to determine if academic achievement in the Five Subjects was different comparing the schools' boarding status at Class 8 where 1=Day school ($n$=7915, $z$=-0.17), 2=Boarding school, ($n$=395, $z$=1.25) and 3=Day and boarding school ($n$=810, $z$=1.00). There was a statistically significant difference between the groups as determined by the one-way ANOVA, $F$(9119) = 1013.48, $p$ <0.001). The Bonferroni post-hoc test showed that the z-score unit difference between boarding and day schools was 1.41, $p$ <.001, while that between Mixed day/ boarding and day schools was 1.17, $p$ <.001. The difference between mixed day/ boarding and full boarding schools was -0.25, $p$ <.001. The effect size, $\eta^2$= 0.18 (measured using eta-squared), was considered large [38].

## 3.3. The Three-Level Random Slope School Model

Table 6 presents the HLM results. The effect of individual predictors remained pretty much the same, and statistically significant, across the three levels as well as through the Five Models.

**Table 6. Three Level Random Slope School Model (Level-3)**

*Fixed Effect*

| Variable | Variable label | Model 1 (Student) | | Model 2 (Teacher) | | Model 3 (School) | | Model 4 (School [Random slope on s21, s61x omitting t22a, t214 and t227]) | | Model 5 (School [Random slope on s21, s61x including t22a, t214 and t227]) | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Est. (Std. Err.) | *p*-value | Est. (Std. Err.) | *p*-value | Est. (Std. Err.) | *p*-value | Est. (Std. Err.) | *p*-value | Est. (Std. Err.) | *p*-value |
| s21 | Female student | -0.27 (0.01) | <.001 | -0.26 (0.01) | <.001 | -0.26 (0.01) | <.001 | -0.25 (0.03) | <.001 | -0.25 (0.03) | <.001 |
| s22a | Student's age in years | -0.08 (0.01) | <.001 | -0.08 (.01) | <.001 | -0.09 (0.01) | <.001 | -0.09 (0.01) | <.001 | -0.09 (0.01) | <.001 |
| s23a | Student's years in current school | -0.02 (0.003) | <.001 | -0.02 (0.003) | <.001 | -0.02 (0.003) | <.001 | -0.02 (0.003) | <.001 | -0.02 (0.003) | <.001 |
| s27 | # of times student spoke English in last 7 days | 0.03 (0.004) | <.001 | 0.03 (0.004) | <.001 | 0.03 (0.004) | <.001 | 0.03 (0.004) | <.001 | 0.05 (0.004) | <.001 |
| s313x | Student's Wealth Index (3 Tertiles): 1=High tertile (Ref.) | | | | | | | | | | |
| | 2=Middle tertile | -0.04 (0.02) | 0.011 | -0.04 (0.02) | 0.012 | -0.04 (0.02) | 0.017 | -0.04 (0.02) | 0.024 | -0.04 (0.02) | 0.023 |
| | 3=Low tertile | -0.13 (0.02) | <.001 | -0.13 (0.02) | <.001 | -0.13 (0.02) | <.001 | -0.12 (0.02) | <.001 | -0.12 (0.02) | <.001 |
| s332 | Mother has some primary education | 0.06 (0.02) | 0.001 | 0.06 (0.02) | 0.001 | 0.06 (0.02) | 0.001 | 0.08 (0.02) | <.001 | 0.08 (0.02) | <.001 |
| s333 | Mother has completed primary education | 0.07 (0.02) | <.001 | 0.07 (0.02) | <.001 | 0.07 (0.02) | <.001 | 0.07 (0.02) | <.001 | 0.07 (0.02) | <.001 |
| s334 | Mother has some secondary education | 0.06 (0.02) | 0.006 | 0.06 (0.02) | 0.006 | 0.06 (0.02) | 0.009 | 0.04 (0.02) | 0.098 | 0.04 (0.024) | 0.097 |
| s353 | Father has completed primary education | -0.07 (0.01) | <.001 | -0.07 (0.02) | <.001 | -0.07 (0.02) | <.001 | -0.07 (0.02) | <.001 | -0.07 (0.02 | <.001 |
| s356 | Father has completed post secondary training | 0.06 (0.02) | 0.002 | 0.06 (0.02) | 0.002 | 0.06 (0.02) | 0.002 | 0.06 (0.02) | 0.003 | 0.06 (0.02) | 0.003 |
| s36c | # of siblings | -0.01 (0.003) | <.001 | -0.01 (0.003) | <.001 | -0.01 (0.003) | <.001 | -0.02 (0.003) | <.001 | -0.02 (0.003) | <.001 |
| s53 | How often learners are hurt by teachers: 1=Never (Ref.) | | | | | | | | | | |
| | 2=Sometimes | -0.05 (0.02) | 0.001 | -0.05 (0.02) | 0.001 | -0.05 (0.02) | 0.001 | -0.06 (0.02) | <.001 | -0.06 (0.02) | <.001 |
| | 3=Often | -0.08 (0.03) | 0.003 | -0.08 (0.03) | 0.003 | -0.08 (0.03) | 0.003 | -0.09 (0.03) | 0.001 | -0.09 (0.08) | 0.001 |
| s58 | Number of times student has repeated classes | -0.08 (0.01) | <.001 | -0.08 (0.01) | <.001 | -0.08 (0.01) | <.001 | -0.09 (0.01) | <.001 | -0.09 (0.01) | <.001 |
| s61x | Student keeps negative company (standardized score) | -0.19 (0.01) | <.001 | -0.19 (0.01) | <.001 | -0.19 (0.01) | <.001 | -0.14 (0.03) | <.001 | -0.14 (0.05) | <.001 |
| t22a | Teacher's age in years | | | -0.01 (0.002) | <.001 | -0.01 (0.002) | 0.001 | omitted | omitted | -0.01 (0.002) | 0.002 |
| t214 | # Number of in-service courses attended (2013) | | | 0.03 (0.01) | 0.029 | 0.03 (0.01) | 0.055 | omitted | omitted | 0.03 (0.01) | 0.05 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| t227 | # of formal written tests in teacher's subject | | | 0.01 (0.003) | 0.001 | 0.01 (0.003) | 0.001 | omitted | omitted | 0.01 (0.003) | 0.001 |
| h16 | Kuria East Sub-County | | | | | -0.85 (0.09) | <.001 | -0.83 (0.09) | <.001 | -0.80 (0.08) | <.001 |
| h24a | Boarding status at class 8: 1=Day (Ref.) | | | | | | | | <.001 | | |
| | 2=Boarding | | | | | 0.98 (0.21) | <.001 | 0.98 (0.21) | <.001 | 0.92 (0.21) | <.001 |
| | 3=Day and boarding | | | | | 0.78 (0.14) | <.001 | 0.83 (0.14) | <.001 | 0.76 (0.13) | <.001 |
| h5122 | School disallows borrowing of library books to take home | | | | | -0.13 (0.02) | <.001 | -0.13 (0.02) | <.001 | -0.13 (0.02) | <.001 |
| | Intercept | 1.58 (0.13) | <.001 | 1.69 (0.14) | <.001 | 1.98 (0.14) | <.001 | 1.93 (0.11) | <.001 | 2.00 (0.14) | <.001 |
| | | *Variance Component* | | *Variance Component* | | *Variance Component* | | *Variance Component* | | *Variance Component* | |
| Student (Level-1), $e_{ijk}$ | | 0.3833 (0.01) | | 0.3832 (0.01) | | 0.3818 (0.01) | | 0.3694 (0.01) | | 0.3694 (0.01) | |
| Teacher (Level-2), $u_{jk}$ | | 0.0509 (0.01) | | 0.0465 (0.01) | | 0.0465 (0.01) | | 0.0514 (0.01) | | 0.0469 (0.01) | |
| School (Level-3), $v_k$ | | 0.4544 (0.08) | | 0.3974 (0.08) | | 0.3818 (0.01) | | 0.1113 (0.02) | | 0.0998 (0.02) | |
| *Variance Explained (%)* | | | | | | | | | | | |
| Student (Level-1), $\sigma^2_e$ | | 0.0529 | | 0.0530 | | 0.0544 | | 0.0669 | | 0.0669 | |
| Teacher (Level-2), $\sigma^2_u$ | | -0.0019 | | 0.0025 | | 0.0025 | | -0.0023 | | 0.0021 | |
| School (Level-3), $\sigma^2_v$ | | 0.0544 | | 0.1118 | | 0.4043 | | 0.3999 | | 0.4114 | |
| | *Variance-Covariance Matrix* | | | | | | | Est. (Std. Err.) | 95% CI | Est. (Std. Err.) | 95% CI |
| | s21, $\sigma^2_{v1}$ | | | | | | | 0.04 (0.01) | 0.02, 0.07 | 0.04 (0.01) | 0.02, 0.07 |
| | s61x, $\sigma^2_{v14}$ | | | | | | | 0.02 (0.01) | 0.01, 0.04 | 0.02 (0.01) | 0.01, 0.04 |
| | Intercept, $\sigma^2_{v0}$ | | | | | | | 0.11 (0.02) | 0.07, 0.17 | 0.10 (0.02) | 0.07, 0.15 |
| | s21,s61x, $\sigma_{v114}$ | | | | | | | 0.00 (0.01) | -0.01, 0.01 | 0.00 (0.01) | -0.01, 0.01 |
| | s21, Intercept, $\sigma_{v01}$ | | | | | | | 0.00 (0.01) | -0.02, 0.02 | -0.00 (0.01) | -0.02, 0.02 |
| | s61x, Intercept, $\sigma_{v014}$ | | | | | | | 0.02 (0.01) | 0.00 (0.04) | 0.02 (0.01) | -0.00, 0.04 |
| *Model Fit Statistics* | | | | | | | | | | | |
| Deviance | | 17824 | | 17797 | | 17690 | | 17565 | | 17541 | |
| Akaike Information Criterion (AIC) | | 17864 | | 17843 | | 17744 | | 17623 | | 17605 | |
| Bayesian Information Criterion (BIC) | | 18006 | | 18006 | | 17936 | | 17829 | | 17833 | |
| Likelihood Ratio test vs. OLS Regression | | $\chi^2$ (2) = 5696 | <.001 | $\chi^2$ (2) = 4815 | <.001 | $\chi^2$ (2) = 2214 | <.001 | $\chi^2$ (7) = 2659 | <.001 | $\chi^2$ (7) = 2363 | <.001 |
| Likelihood Ratio test (Preceding vs. Current Model) | | $\chi^2$ (16) = 1140 | <.001 | $\chi^2$ (3) = 26 | <.001 | $\chi^2$ (6) = 107 | <.001 | na | na | $\chi^2$ (3) = 24 | <.001 |

*Note. N*= 9120 (1824*5 = 9120, each student has 5 records); Est. = Estimate; Std. Err. = Standard Error (in parentheses); AIC and BIC statistics = smaller-is-better fit; OLS=Ordinary Least Squares; CI= Confidence Interval; na=not applicable; #=Number

The results are now discussed in the following three sub-sections. The results of variance partitioning and variance explained are also discussed in each sub-section.

### 3.3.1. Student-Level Predictors

All 14 predictors at student-level had significant influence on the outcome variable. Students' sex (s21), their wealth index (s313x) and keeping of negative company (s61x) had the largest standardized regression coefficients of ≥0.10 [5,39] and were considered to have the greatest influence on the outcome variable at student-level. Female students were estimated to score up to -0.25 standard deviation units below what a male student with similar characteristics would. Despite affirmative action programmes in Kenya targeting female student, this disadvantage seems to persist in the literature with these results confirming those found by other researchers [5,6].

Students from the middle and low wealth index tertile performed lower than their counterparts in the high tertile. This appears to be a "double tragedy" that those disadvantaged in wealth are also disadvantaged in academic achievement [5,6,40,41]. Most of the students with high scores were either in top of the range private schools or in boarding schools. Most of these schools are associated with wealthier households that can afford high school fees and additional levies for improvement of the learning environment. Education in public primary schools in Kenya is subsidized by the Government through tuition fees with parents shouldering the costs of school meals (in non school meal programme schools), uniform and transport among other costs. We argue that there are public schools that perform equally well and that it is possible to improve their teaching-learning processes and environment. Many of these public schools are often in rural areas where most of the less wealthy households live.

Students who kept negative company (s61x) scored lower than those who did not. A one standard deviation increase in keeping negative company was associated with up to -0.14 (*SE*=0.05) standard deviation units below the mean. Negative behaviour was defined as company with friends who took alcohol and/or sneaked away from home without parental permission and/or got/get into trouble with school administration or Police for something bad they did and/or engaged in sex or sexual activity and/or smoked cigarettes or used hard drugs such as bang' and/or got into fights and quarrels with other people. Positive behaviour was defined as regular attendance of church/mosque and/or desire to join secondary school, and/or working hard in academic work, and/or get good marks in academic work and/or been commended or given a gift for good work or good behaviour. This result is expected as negative behaviour distracts a student's focus on academic pursuit [42,43].

When first modelled at Leve-1, student-level predictors explained up to 0.1053 (10.53%) of the variation in student scores across the three levels (Level-1, 5.29%, Level-2, -0.29% and Level-3, 5.44%). Focussing only at Level-1 while adjusting for predictors at levels 2 and 3, student-level predictors explained 6.69% of the 43.88% variance available for explanation in the unconditional model in Table 4 leaving up to 37.19% of the variation unexplained. This suggests that future research on student-level variables probably needs to go beyond those modelled in this paper.

### 3.3.2. Teacher-Level Predictors

This paper focussed on the effect of teacher-level variables on student academic achievement. An extra year in teacher-age impacted negatively on student scores by 0.01 (*SE*=0.002) standard deviation units below the mean. This result suggests a "double edged sword" because it is naturally expected that as teachers advance in age, they probably become more experienced and are better placed to benefit students more compared with newly employed graduates from college. These results do not support this argument. Instead, ageing comes with natural "tare and ware" and tiredness with teachers beginning to think and plan for their retirement from service than thinking of how to improve learning outcomes. It is common practice to find that "retiring" teachers in public schools are given "less demanding classes" (and there are no such Classes as all are demanding at their respective level), often at lower primary (Classes 1-3) or at Kindergarten. These results paint a grim picture going forward because the Teachers Service Commission and Ministry of Education adjusted the teachers' retirement age upwards to 60 from 55 which means that teachers now retire when much older than before.

Teachers who attended short in-service courses in their respective Subject areas had a positive impact on their students' KCPE scores. An extra such course attended was associated with an increase of 0.03 (*SE*=0.01) standard deviation units above the mean. The number of formal written tests in a teacher's Subject area also had a positive impact on student scores. We note that there is currently a lot of testing in all Classes. This ranges from once weekly to once monthly with the frequency and intensity increasing at Classes 7 and 8. Sadly, this often compromises normal teaching and exposure to syllabi content with the focus shifting to exposure to "KCPE- examination-like" tests than content coverage. The authors caution that exposure to such tests in preparation for the final examination is good for the learners' revision but should be done with moderation with syllabi content given first priority.

Adjusting for predictors at Leve-1, the effect of the three teacher variables was substantial across the three levels. When introduced in Model-2, they marginally improved the variance explained at Level-1 from 5.29% to 5.30% and accounted for 0.44% at their own Level-2 (Model-2, 0.0025 minus Model-1, -0.0019). But their effect was felt more at Level-3 whose proportion of explained variation improved from 0.0544 to 0.1118 (a difference of 0.0574 or 5.74%). So the net value of the three teacher predictors across the three levels was 6.19% (0.01+0.44+5.74). Adjusting for predictors at levels 1 and 3 in the final Model-5, the variance explained by the three teacher predictors at Level-2 was 0.21% of the 4.93% available for explanation at that level in the unconditional model in Table 4. A lot more teacher variables had been considered with up to 15 returning no statistically significant correlation with the outcome variable (see Table 5). With the three teacher variables not meeting the standardised regression coefficient of ≥0.10 to be flagged as predictors of student academic achievement (though statistically significant), these results support the findings of researchers who have argued that teacher or school resource inputs do not explain a large portion of student academic achievement. For instance, Hanushek [20,21,44] reviewed up to 400 studies on expenditure relationships in schools and argued in his conclusion that there was no strong evidence that pupil teacher ratio, teacher education, or teacher experience had the expected positive effects on student achievement and that there was no strong or systematic relationship between school expenditures and student performance. Other researchers have however found substantial effect, see for example [7,8].

### 3.3.3. School-Level Predictors

The three school-level predictors had huge influence on student academic achievement. Schools in Kuria East Sub-County scored up to negative 0.80 (SE=0.08) standard deviation units below what schools in Mumias Sub-County would score. This confirms Kuria East's low ranking on the merit list for the period 2010-2012. Mumias Sub-County was ranked in the top 5% percent Sub-Counties for the same period. Boarding schools also had a large impact with close to one standard deviation above what day schools would score. A boarding component within a day school (often involving Classes 6-8) also had a positive effect compared with day schools. Boarding schools certainly have an advantage of "more time for students" over day schools whose students often have to attend to home chores after school. This eats into their time for private study and completion of homework. There is a move in Kenya towards establishing more day schools at secondary school because they are cheaper to run compared with boarding schools which are also prone to indiscipline. For instance, within second term (May-August 2016), more than 100 boarding secondary schools in Kenya had reported fire incidences with dormitories or other infrastructure being torched by students for a myriad of reasons that are still under investigation. No boarding primary school had reported any such cases but this is an indication of the "soft underbelly" that boarding schools are.

Adjusting for predictors at levels 1 and 2, school-level predictors explained 41.14% of the 48.04% variance explained by all the levels in the Model-5. This means that this level accounted for a large proportion of the variation seen in student academic achievement in the KCPE examination. This is within the range found by other studies of near-similar grade levels in Kenya [5] and other developing countries. For instance, in Latin American countries, the variance between schools in mathematics achievement among Grades 3 and 5 pupils ranged from 19.5% to 41.2% [45].

### 3.3.4. The Random Slopes Model

Two student-level covariates, sex of student as well as student keeps negative company (z-score) had the largest effects on student academic achievement in Model-3 with the model assuming that the variability in student-specific deviations from the intercept was the same for female and male students as well as same for students who kept positive or negative company. To check this assumption, these two covariates were introduced into the random component of the model. Model-5 (the random slopes model) allowed a random intercept as well as random slopes. The fixed effect referred to the overall expected effect of a student's gender on KCPE academic subject-scores while the random effect gave information on whether or not this effect differed across schools. Therefore, subtracting the variance explained in Model-3 (46.12%) from that explained in Model-5 (48.04%), we get the net-variance effect of the slopes model as 1.92%, confirming that the effect of the two covariates differed across schools.

## 4. Conclusion and Implications for Policy

With ageing teachers having negative effect on student scores, the policy that adjusted teacher retirement age from 55 to 60 years should probably be reviewed. The policy seems to have been driven more by considerations surrounding the payment of retirement-packages for teachers and had probably nothing to do with improving student scores if those teachers stayed on till age 60. This policy shift was not backed by a large body of credible research evidence. An extra short-term in-service course attended by teachers in their respective subject areas had positive effect on student scores in those subjects. This result points to probable positive dividends for students and schools that could be reaped from continued teacher-learning beyond college. Stake holders in school management should probably be encouraged to avail resources to enable teachers improve their teaching. “Scarcity” of resources for such short in-service courses is often the excuse used to deny teachers participation. School administrators and their Boards of Management should therefore be encouraged to put premium on what improves learning outcomes. While exposure to more testing is beneficial in improving student scores, this should be balanced with equal effort at syllabus coverage and student exposure to subject-specific content that goes beyond putting a premium on testing per se.

## Statement of Competing Interests

The authors have no competing interests.

## References


[1] Coleman, J.S., et al., *Equality of Education Opportunity (Coleman) Study (EEOS)*. 1966, Washington, DC: Department of Health, Education, and Welfare, Office of Education/National Center for Education Statistics.

[2] Hanushek, E.A. and S.G. Rivkin, *Harming the Best: How Schools Affect the Black–White Achievement Gap.* Journal of Policy Analysis and Management, 2009. 28(3): p. 366-393.

[3] Carbonaro, W. and C. Elizabeth, *School Sector and Student Achievement in the Era of Standards Based Reforms.* Sociology of Education, 2010. 83(2): p. 160-182.

[4] Southern and Eastern Africa Consortium for Measuring Education Quality, *Quality of primary school inputs in Kenya*. 2011, SACMEQ: Harare, Zimbabwe.

[5] Hungi, N. and F.W. Thuku, *Differences in pupil achievement in Kenya: Implications for policy and practice.* International Journal of Educational Development, 2010. 30(1): p. 33-43.

[6] Ejakait, E., et al., *Factors associated with low achievement among students from Nairobi's urban informal neighborhoods.* Urban Education, 2011. 46(5): p. 1056-1077.

[7] Neild, R.C., *The Effect of Teacher Certification on Middle Grades Achievement in an Urban District.* Educational Policy, 2009. 23(5): p. 732-760.

[8] Konstantopoulos, S. and V. Chung, *The persistence of teacher effects in elementary grades.* American Educational Research Journal, 2011. 48(2): p. 361-386.

[9] Legewie, J. and T.A. DiPrete, *School Context and the Gender Gap in Educational Achievement.* American Sociological Review, 2012. 77(3): p. 463-485.

[10] Heyneman, S.P. and W.A. Loxley, *Influences on Academic Achievement Across High and Low Income Countries: A Re-Analysis of IEA Data.* Sociology of Education, 1982. 55(1): p. 13-21.

[11] Heyneman, S.P. and W.A. Loxley, *The Effect of Primary-School Quality on Academic Achievement Across Twenty-nine High and Low-Income Countries.* The American Journal of Sociology, 1983. 88(6): p. 1162-1194.

[12] Sanders, W.L. and J.C. Rivers, *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*. 1996, University of Tennessee Value-Added Research and Assessment Center: Knoxville, Tennessee.

[13] Wright, S.P., S.P. Horn, and W.L. Sanders, *Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation.* Journal of Personnel Evaluation in Education, 1997. 11(1): p. 57-67.

[14] Darling-Hammond, L., *Teacher Quality and Student Achievement: A Review of State Policy Evidence*. 1999, Center for the Study of Teaching and Policy, University of Washington: Seattle, Washington.

[15] Laczko-Kerr, I. and D.C. Berliner, *The Effectiveness of "Teach for America" and Other Under-certified Teachers on Student Academic Achievement: A Case of Harmful Public Policy.* Education Policy Analysis Archives, 2002. 10(37): p. 1-53.

[16] Wenglinsky, H., *How Schools Matter: The Link Between Teacher Classroom Practices and Student Academic Performance.* Education Policy Analysis Archives, 2002. 10(12): p. 1-30.

[17] Nye, B., L.V. Hedges, and S. Konstantopoulos, *Do Low-Achieving Students Benefit More from Small Classes? Evidence from the Tennessee Class Size Experiment.* Educational Evaluation and Policy Analysis, 2002. 24(3): p. 201-217.

[18] Rivkin, S.G., E.A. Hanushek, and J.F. Kain, *Teachers, Schools, and Academic Achievement.* Econometrica, 2005. 73(2): p. 417-458.

[19] Caprara, G.V., et al., *Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level.* Journal of School Psychology, 2006. 44(6): p. 473-490.

[20] Hanushek, E.A., *The Economics of Schooling: Production and Efficiency in Public Schools.* Journal of Economic Literature, 1986. 24(3): p. 1141-1177.

[21] Hanushek, E.A., *The impact of differential expenditures on school performance.* Educational Researcher, 1989. 18(4): p. 45.

[22] Hanushek, E.A., *Assessing the Effects of School Resources on Student Performance: An Update.* Educational Evaluation and Policy Analysis, 1997. 19(2): p. 141-164.

[23] Rabe-Hesketh, S. and A. Skrondal, *Multilevel and Longitudinal Modeling Using Stata, Volumes I and II*. Third ed. 2012, College Station, Texas USA: Stata Press.

[24] Bryk, A.S. and S.W. Raudenbush, *Hierarchichal Linear Models: Applications and Data Analysis Methods*. 1992, Newbury Park, California USA: SAGE Publications, Inc.

[25] Raudenbush, S.W. and A.S. Bryk, *Hierarchical Linear Models: Application and Data Analysis Methods*. Second ed. 2002, Thousand Oaks, California: Sage Publications, Inc.

[26] Goldstein, H., *Multilevel statistical models*. 2nd ed. 1995, New York, USA: Halsted Press.

[27] Hardin, J.W. and J.M. Hilbe, *Generalized Linear Models and Extensions*. 3rd ed. 2012, College Station, Texas USA: Stata Press.

[28] Butt, J.J., *The Greenwood Dictionary of World History*. 2006, Westport, USA: Greenwood Press.

[29] Chiraghdin, S. and M.E. Mnyampala, *Historia Ya Kiswahili*. 1977, Nairobi: Oxford University Press.

[30] Massamba, D.P.B., *Historia Ya Kiswahili*. 2002, Nairobi: Jomo Kenyatta Foundation.

[31] Leckie, G., *Three-Level Multilevel Models*. 2013, Bristol, UK: Centre For Multilevel Modelling, University of Bristol.

[32] Leckie, G., *Introduction to Multilevel Modelling*. 2010, Bristol, UK: Centre For Multilevel Modelling, University of Bristol.

[33] StataCorp (2013) *Stata Multilevel Mixed-Effects Reference Manual: Release 13*.

[34] Filmer, D. and L.H. Pritchett, *Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrolments in States of India.* Demography, 2001. 38(1): p. 115-132.

[35] Rutstein, S.O. and J. Kiersten, *The DHS Wealth Index*. 2004, Calverton, Maryland USA: ORC Macro.

[36] ORC Macro. *The DHS Program, Demographic and Health Surveys: Wealth Index*. 2016 [cited 2016 18/07]; Available from: http://www.dhsprogram.com/topics/wealth-index/Index.cfm.

[37] Taylor, R., *Interpretation of correlation coefficient.* Journal of Diagnostics Medical Sonography, 1990. 6(1): p. 35-39.

[38] Acock, C.A., *A Gentle Introduction to Stata*. 2006, College Station, Texas USA: Stata Press.

[39] Hox, J.J., *Applied Multilevel Analysis*. 1995, Amsterdam, The Netherlands: TT-Publikaties.

[40] Daniel, H.C., *Socioeconomic Status and Academic Achievement: Trajectories from Childhood to Adolescence.* Canadian Journal of Education, 2009. 32(3): p. 558-590.

[41] Jensen, E., *Teaching with Poverty in Mind: What Being Poor Does to Kids' Brains and What Schools Can Do About It*. 2009, Alexandria, VA USA: ASCD.

[42] Gutman, L.M. and J. Vorhaus, *The Impact of PupilBehaviour and Wellbeing on Educational Outcomes*. 2012, London, UK: Childhood Wellbeing Research Centre, Institute of Education, University of London.

[43] Valiente, C., J. Swanson, and N. Eisenberg, *Linking Students' Emotions and Academic Achievement: When and Why Emotions Matter.* Child Dev Perspect 2012. 6(2): p. 129-135.

[44] Hanushek, E.A., *Hanushek Responds.* Educational Researcher, 1989. 18(6): p. 25.

[45] Willms, D.J. and M.A. Somers, *Family, classroom, and school effects on children's educational outcomes in Latin America.* School Effectiveness and School Improvement, 2001. 12(4): p. 409-445.