Faculty of Mathematics and Informatics

VILNIUS UNIVERSITY

FACULTY OF MATHEMATICS AND INFORMATICS
MASTER'S STUDY PROGRAMME
MODELLING AND DATA ANALYSIS

Student Performance from a Socio-Economic Perspective

Master's Thesis

Author: Paulius Brandišauskas

VU email address: paulius.brandisauskas@mif.stud.vu.lt

Supervisor: prof. hab. dr. Vydas Čekanavičius

Vilnius
2023

Mokinių pasiekimai vertinant iš socialinės ir ekonominės perspektyvos

Santrauka

Šiame tyrime yra nagrinėjamas ryšys tarp gamtos mokslų testo rezultatų ir socialinės bei ekonominės aplinkos mokinių ir mokyklų lygmeniu Lietuvoje. Analizuojant 2015 m. PISA duomenis ir naudojant naujausią R programinės įrangos paketą, yra atliekamas hierarchinis tiesinis modeliavimas su daugiapakopiais imties svoriais ir tikėtinomis reikšmėmis. Siekiant nustatyti, kurie socialiniai ir ekonominiai veiksniai yra reikšmingi mokinių pasiekimams, pateikiami du atskiri galutiniai devintos ir dešimtos klasės mokinių modeliai. Gauti rezultatai galėtų būti aktualūs Lietuvos švietimo bendruomenei, siekiant suprasti skirtumus tarp vaikų iš skirtingų socialinių sluoksnių. Gilesnis skirtingų pasiekimų priežasčių suvokimas galėtų turėti reikšmės užtikrinant socialinį teisingumą ir padedant mažiau galimybių turintiems vaikams pasiekti daugiau.

Raktiniai žodžiai: Gamtamokslinis raštingumas, hierarchinis tiesinis modelis, PISA duomenų analizė, socialinis ir ekonominis statusas

Student Performance from a Socio-Economic Perspective

Abstract

This study examines the relationship between the science test results and the socio-economic environment at student and school levels in Lithuania. Hierarchical linear modelling with multilevel sample weights, plausible values and the latest R software package are used to analyse the Programme for International Student Assessment 2015 data. Two separate final models are presented for ninth and tenth graders to identify which socio-economic factors are significant for student achievement. The results could be relevant for the Lithuanian education community to understand the differences between children from different backgrounds. A deeper perception of the causes of differences in performance could have implications for facilitating social justice and helping children with fewer opportunities to achieve more.

Keywords: Scientific literacy, hierarchical linear modelling, PISA data analysis, socio-economic status

Contents

1 Introduction

Lithuania started participating in PISA assessment in 2006. Although Lithuania is one of the fastest developing countries in the OECD, its science literacy rate is still below the OECD average in all student testing periods from 2006 till 2018. There is also no clear upward trend in performance across the board, as Lithuania's average science literacy score was 488 in 2006, 491 in 2009 and 496 points in 2012. In 2015, a computer-based assessment (CBA) was introduced and the majority of the countries moved from a paper-based assessment to a new method and Lithuanian students achievement growth has stopped since Lithuania scored only 475 in 2015 and 482 points in 2018. Such a pronounced achievement gap is unlikely to be explained by the new testing method, as there are no such noticeable drops in the other testing domains in the period of 2012 to 2015.

Lithuanian education system is not at its best. According to Lithuanian president G. Nausėda the main problems include social exclusion in education, underfunding of research, the quality of educational institutions, and problems with teacher training [13]. Nerija Putinaitė, Associate Professor at the Institute of International Relations and Political Science at Vilnius University and former Deputy Minister of Education and Science, agrees that one of the main problems is the gap in mainstream schools, the so-called quality scissors [13]. In another interview, she also expresses that rural schools cannot produce good results no matter how many teachers are employed – children from socially at-risk families need to be among motivated students to get different socialization skills [8]. Indeed, the socio-economic background is very important for student achievement. According to Valdemaras Razumas, Deputy Minister of Education, Science and Sport, says the situation is dreadful. He claims that one of the main reasons for this problem is the social divide between students studying in the regions and those studying in the cities. V. Razumas adds that the network of smaller urban and rural schools needs to be optimized [8].

Analysis of results from international surveys (e.g., TIMMS, PISA) in different countries shows that, regardless of the country, socio-economic and cultural exclusion is clearly linked to student achievement. This means that all countries face challenges in terms of ensuring equality of opportunity and effectiveness in education and training. Nonetheless, some students achieve high academic results despite their socio-economic disadvantages. On average in OECD countries, one in ten disadvantaged students can show very high reading achievement and be in the top quarter of achievers. This shows that disadvantage is not fatal for a student. Moreover, in Australia, Canada, Estonia, Hong Kong (China), Ireland, Macao (China) and the United Kingdom where all students performed above the OECD average, more than 13% of disadvantaged students were academically resilient, i.e. in the overall top quarter of attainment [24]. In Lithuania, some of the most famous studies regarding student achievement and their background have been carried out by A. Zabulionis [32] and A. Jaržemskis et al. [9].

This study uses publicly available data from PISA 2015, when science literacy was major research domain, to obtain significant regressors that affect student achievement and to investigate the impact of not only the most studied variables (students' socio-economic status or the location of the school) on students' performance but also less examined – use of digital devices at school, proportion of teachers with particular education, etc. To perform the research hierarchical linear modelling is employed with multilevel weights. The methodology of hierarchical linear modelling and data used for models are described in more details in sections 3 and 4, respectively. In section 5 explanatory analysis is presented for Lithuanian dataset as well as an initial multilevel pooled model and logic behind it. Final models presented in Result section for grades 9 and 10 (denoted accordingly as Equations 4 and 5) are obtained by removing insignificant variables and taking into account correlation matrix. The findings presented in Result section could potentially provide insights into what more could be done in the much-needed education reform in Lithuania.

2 Literature review

The relationship between socio-economic status (SES) and student achievement is a frequently studied topic in an academic world. SES is often considered to be one of the key student and school factors predicting student achievement. Research shows that students from lower SES backgrounds tend to have worse achievements. Therefore, solutions are being sought at both school and education system level to help these students to learn and achieve at higher degrees.

Most famous associations gathering statistical data about scientifical literacy and student living environment are International Association for the Evaluation of Educational Achievement (IEA) and The Organisation for Economic Co-operation and Development (OECD). IEA large-scale assessment called Trends in International Mathematics and Science Study (TIMSS) and OECD initiative Programme for International Student Assessment (PISA) involves more than 40 countries and thousands of students in each participating country. Numerous linked research initiatives and a large number of publications resulted from these studies. While TIMSS collects data on educational achievement from students in fourth and eighth grades and puts formal scientific knowledge taught in school to test, PISA assessments are administered to mainly 15-years-old students. The tests are supposed to assess how well students are prepared for real life situations waiting for them in the future. The literacy idea places a strong emphasis on process mastery, conceptual comprehension, knowledge application, and functioning in a variety of contexts. PISA's emphasis on scientific literacy allows it to draw from both classroom curricula and potential extracurricular learning when TIMSS focuses only on school syllabus and educational systems of different nations vary, therefore PISA survey is used to administer the performance from different perspectives. Since PISA intends to assess alleged 'life skills' gather data to guide educational decisions, and track the effectiveness of the educational system, PISA is more criticized than TIMSS because of its more ambitious goals [1]. Despite the controversy, the number of countries participating in PISA has increased over the years. Standardized exams, such as PISA, help the researchers to make sense of a complicated world by supplying data that enables comparisons both within and across nations. PISA increased awareness of aspects besides the quantity of time spent in a classroom that influence students' achievements. Unquestionably, this was a step forward for comparative education research [1].

As mentioned above, the relationship between SES and academic achievement has attracted a lot of scientific attention. Unfortunately, no agreement was reached on the conceptual meaning of socio-economic status or how to assess it in studies looking at how it links to educational attainment and achievement among school-aged children. To define social class, poverty and wealthiness, or a person's standing on the social hierarchy, many variables, or combinations of factors, are frequently used interchangeably [2, 28, 30]. For example, index of economic, social, and cultural status (ESCS), a composite measure that compiles the financial, social, cultural, and human capital resources accessible to students into a single score, is used in PISA to determine a student's socio-economic position. In reality, it is derived from a number of factors pertaining to students' family backgrounds that are then divided into three categories: parents' educational attainment, parents' occupations, and an index summarizing a number of home possessions that can be used as proxies for material wealth or cultural capital, such as owning a car, having a quiet workspace, having access to the internet, and the number of books and other educational resources in the home [25]. Academic success and SES of individual students are positively correlated. Strong and beneficial evidence supports this link; usually, superior educational results are connected with higher student level SES [11, 14, 15]. Moreover, for kids from various socio-economic backgrounds, previous studies have looked at differences in the relationship between school composition and success. For instance, five decades ago Coleman discovered that

African-American students profited from attending a school with children who were from higher socio-economic levels [5]. More recent research has revealed a high correlation between success and school SES for all children [3, 15, 31]. It was shown that students who had peers with higher SES levels fared better academically, perhaps as a result of the more resources provided to students at schools with higher SES populations [4]. Therefore, student success rises with economic prosperity, while economic prosperity rises with increasing levels of education. In addition to being fair, reducing socio-economic school segregation is also productive. For instance, Canada and Finland do better on PISA than Australia, yet there is a less correlation between school SES and student achievement in those two nations than there is in Australia [15]. As these nations demonstrate, removing socio-economic barriers between schools and within schools encourages greater overall student accomplishment for all children without affecting the achievement of high-achieving kids [27]. PISA report finds that the highest performing nations educate all their students to a high level, not just some of them [17]. Although the relationships between students' performance and SES are well studied on both student and school levels, less is known about the situation when both levels are included into the equation.

In Lithuania, A. Zabulionis [32] conducted a study of PISA 2018 data of Lithuanian students achievement. According to the author, students' academic performance depends on two interrelated factors - the student's SES and the student's school location. The achievement gap between urban and rural schools is widening and the Lithuanian education system seems to face the greatest challenges in ensuring equal education opportunities for all students in both rural and urban schools. One of the most important attributes of the quality of an education system should be the provision of equal access to education, which requires equal opportunities for students with similar SES status, regardless of their school location. Lithuanian results were also studied by A. Jaržemskis et al. [9]. The research included 8th graders TIMSS data and in all cycles (2003, 2007, 2011, 2015 and 2019) it has been found that the more students in a school, living in a high-SES context, the higher math and science achievements are and vice versa. Also, for both math and science, a stronger correlation was found between the students living in low-SES contexts and the achievement than students living in a high-SES context. Although these secondary analyses provide an in-depth insight into the state of the Lithuanian education system, they are descriptive and exploratory in nature and do not provide data modelling.

One of the frequently used tools to measure the relationship between student achievement and the environment he/she is living in is hierarchical linear modelling (HLM). Jui-Chen Hsu using HLM and PISA 2003 data, discovered that in Canada and Hong Kong 'school composition has an effect on mathematics achievement over and above that of individual characteristics.'. The student characteristics included SES, sex, family structure and immigration background. Also, the findings revealed that schools in Canada and Hong Kong, each accounted for 20% and 49% of the variance in math achievement, respectively [6]. Haigen Huang, who also used HLM and PISA 2012 data, indicated that students in United States of America who believed they were persistent were more likely to achieve well than those who believed they were less persistent. Additionally, more time spent studying in school was linked to better performance. High-SES students did, however, spend more time in class studying and thought of themselves as more persistent. Therefore, for the majority of low-SES children, learning time and tenacity were not likely to remove the SES limitation on accomplishment unless schools gave them more classes and learning opportunities [7]. Another HLM research conducted in USA, using TIMSS 2003 data, showed that along with the statistically significant impact of the teachers' participation to students' science achievement, other important criteria were the topic coverage and teaching certification in science, the impact of the school's SES, and the availability of remedial and enrichment science programs [12].

3 Methodology

According to J. W. Osborne students are a part of a hierarchical social system that may include their family, classroom, grade level, school, nation etc. Students who are part of hierarchies resemble one another more frequently than students who were randomly selected from the overall population. Students in a specific classroom, for instance, could experience a similar atmosphere, which over time may increases homogeneity [26]. One of frequently used methods called HLM is designed to take into consideration the hierarchical structure of educational data, where hierarchies occur naturally. In this work two-level model is used. The first level is student level where the model estimates how student level predictors impact the outcome variable. At the school level, the predictors estimate how school level variables are related to the average outcome across schools [29].

3.1 Unconditional model

To determine if the hierarchical structure has to be taken into consideration at all, an unconditional (null) model is used. It provides the variability of both levels and indicates if there is significant variation at school level to justify the use of HLM. The unconditional model equations are:

$$\begin{aligned} \text{Student level}: Y_{ij} &= \beta_{0j} + e_{ij}, \quad i = 1, ..., n_j, \quad j = 1, ..., J; \\ \text{School level}: \beta_{0j} &= \gamma_{00} + u_{0j}, \quad j = 1, ..., J. \end{aligned}$$

Where Y_{ij} is the outcome variable for ith person from jth school and $i = 1, ..., n_j$, $j = 1, ..., J$. β_{0j} is called an intercept or a mean of an outcome variable of jth school, e_{ij} is the individual difference from this mean for student i from school j. At school level, γ_{00} is the mean score for all schools (overall intercept) and u_{0j} is considered the difference from that mean or school level error. The error terms e_{ij} and u_{0j} have variances σ^2 and τ_{00}, respectively. To measure the degree of variability between groups (in our case schools) the primary characteristic of the null model called Intraclass Correlation Coefficient (ICC) is used, its formula is:

$$ICC = \frac{\tau_{00}}{\tau_{00} + \sigma^2}. \tag{1}$$

ICC can also be interpreted as a tool to administer how much of the variation between students can be explained by the differences of schools.

3.2 Multilevel model

The PISA data is gathered from a sample of individual students at the chosen schools in the different areas. Because of this aspect, data is very hierarchical, necessitating a multilevel statistical analysis. As mentioned in the Literature review, scientists use HLM to address the multilevel problem. Therefore, the model equations including student and school levels are:

$$
\begin{aligned}
\text{Student level: } Y_{ij} &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + ... + \beta_{Kj}X_{Kij} + e_{ij};\\
\text{School level: } \beta_{0j} &= \gamma_{00} + \gamma_{01}Z_{1j} + ... + \gamma_{0S}Z_{Sj} + u_{0j};\\
\beta_{1j} &= \gamma_{10} + \gamma_{11}Z_{1j} + ... + \gamma_{1S}Z_{Sj} + u_{1j};\\
&\quad \text{...}\\
\beta_{Kj} &= \gamma_{K0} + \gamma_{K1}Z_{1j} + ... + \gamma_{KS}Z_{Sj} + u_{Kj};
\end{aligned}
$$
where $i = 1, ..., n_j, \quad j = 1, ..., J.$

Here X_{Kij} denotes the result (predictor variable) of kth variable for ith respondent from jth school and Z_{Sj} is the predictor variable at school level of sth variable for jth school. β_{Kj} is the slope for the student level predictor variable X_{Kij}. e_{ij} and u_{Kj} are student and school level errors and have variances σ^2 and τ_{KK}, respectively. $\gamma_{10}, ..., \gamma_{K0}$ represent the coefficients for the predictors at student level, while $\gamma_{01}, ..., \gamma_{0S}$ coefficients refer to the student level predictor variables of the school level intercept and estimate the relationship between b_{0j} and $Z_{1j}, ..., Z_{Sj}$. $\gamma_{11}, ..., \gamma_{1S}, \gamma_{K1}, ...\gamma_{Sj}$ are coefficients estimating the association between variables $Z_{1j}, ..., Z_{Sj}$ and slopes $\beta_{1j}, ..., \beta_{Kj}$.

4 Data

The OECD initiative PISA is a cooperative project between OECD Member nations and non-Member partner countries to assess how well 15-year-old students are equipped to face the challenges of today's society. The evaluation is prospective; rather than concentrating on the degree to which these students have mastered a particular school curriculum, it examines their capacity to use their knowledge and abilities to tackle challenging situations in real life. This perspective indicates a shift in curriculum aims and objectives, concentrating more on what students can accomplish with their education. Every three years, the PISA survey is conducted to show students literacy in three domains: mathematics, reading and science. PISA is an age-based survey that evaluates students in grade 7 or above who are 15 years old. In most participating nations and economies, these students are nearing the conclusion of their obligatory education, and enrolment in schools at this level is nearly ubiquitous [20]. Lithuania started to take part in this study in 2006 and in 2015 it was 4th time it took part in PISA examination. The OECD's PISA 2015 is the source of all the data utilized in this study.

4.1 Sampling

Each country's target population for PISA 2015 consisted of 15-year-old students enrolled in schools in grades 7 and above. Typically, surveys in the field of education use a two-stage stratified sample design, therefore, it is not surprising that this design was employed for the particular assessment, meaning:

- First-stage sample units were made up of individual schools with 15-year-old students, or those with the potential to have such children at the time of the evaluation. Schools were systematically selected from a nationwide list of all PISA-eligible schools with probabilities proportional to a measure of size.
- In the second-stage students within sampled schools were chosen with equal probabilities. These students formed a sample, which size was generally 42 for CBA countries. If the PISA-eligible number of students within the school was smaller than 42, all students were selected.

To ensure reliable estimates of student performance, PISA selected the sample of students using developed and widely accepted standards for scientific sampling, in a manner that guaranteed the representation of the whole target population of 15-year-olds. In computer-based nations like Lithuania, a minimum sample size of 5250 assessed students and 150 schools had to be attained. Lithuania has met the minimum requirements, with 6525 students from 311 schools participating.

4.2 PISA test design and context questionnaires

Under the direction of the PISA Governing Board, it was decided to go from a predominantly paper-based delivery survey with optional computer-based modules to a wholly computer-based assignment for the 2015 cycle. Although there was still an option for countries to implement paper-based survey, Lithuanian government chose CBA testing. In 2015 the scientific literacy was the main domain of coverage and in the assessment design for scientific literacy a total of 184 (99 new, 85 trend) items were used. For comparison mathematics and reading total items consisted of 81 and 103 items, respectively. The 2015 assessment tools only included trend items from earlier tests for math and reading. Items related to science included both 2015 trends and brand-new items.

To provide the most thorough assessment of literacy PISA would need to give each student the whole collection of test items in order. The easiest method to close any gaps or biases in the evaluation would be to ask students to respond to all of these questions but the exam would take hours to complete as a result of this. Thus, all PISA cycles, including 2015, had test content that was broken up into many 30-minute clusters or test booklets to make it possible to assess student competency across all subject areas.

The 2015 CBA included three clusters of new collaborative problem-solving materials, six clusters of new science literacy test problems, and six clusters from each of the trend areas of science, reading, and mathematics literacy. Using 66 test forms to divide into six groups, the clusters were distributed in a rotating pattern. Consequently, the assessment was designed such that each student would have two hours to complete the four domains of reading, mathematics, science, and collaborative problem-solving material. Students took an hour of scientific literacy testing in addition to an hour in another subject or two clusters lasting 30 minutes from two of the other three minor domains. Therefore, every student who took the test responded to the questions about science, but not every student responded to the questions on mathematical literacy, reading literacy, or collaborative problem solving [23]. A description summary of the CBA design is shown in Figure 1.

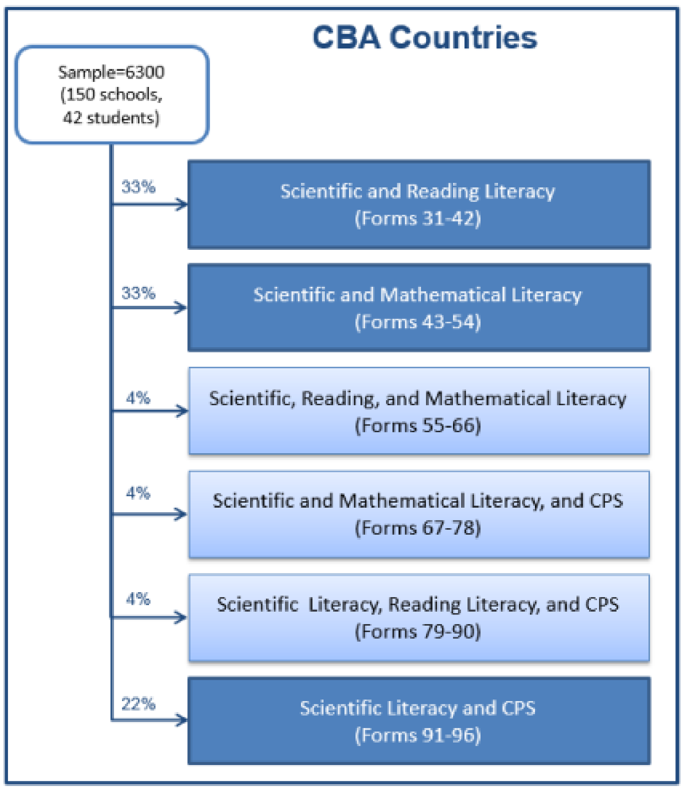

Figure 1: Summary of CBA design [23]

Following the cognitive evaluation, students also answered a 30-minute survey on their background, attitudes, and educational experiences. A 30-minute survey that asked questions about the structure, resources, instruction, culture, and policies of the schools where PISA was given was also completed by the principals. These were so called context questionnaires that give details on the learning environment at the individual, school and national levels of education. Since the initial questionnaire framework for PISA 2009 was published, many methodologies have been used to influence the production of PISA questionnaires. The framework and questionnaire development for PISA 2015 sought to combine the pre-existing approaches with new aspects of policy interest that are currently guiding the discussion on educational effectiveness and education policy decisions. Earlier frameworks focused on the hierarchical structure of educational systems (PISA 2009) and questions of educational effectiveness (PISA 2012). For more information see [18].

4.3 Instruments

Instruments used in this work are reported values from PISA test and context questionnaires. In the subsection part of Instruments the ideas about the plausible values (dependent variables) and regressors for HLM are presented.

4.3.1 Plausible values

Traditionally, when analysing large-scale assessment program (TIMSS, PISA or National Assessment of Educational Progress) results, plausible values (PVs) are reported instead of students' actual scores. The PVs are often viewed as a representation of the variety of competencies each student possesses. The main advantage of PVs is arguably better representation of investigated phenomena. Indeed, in the PISA 2015 test format various student groups responded to different but overlapping sets of items. Usually, it is argued that any statistic based on the quantity of accurate replies should thus not be used when reporting the survey findings because student actual score may not represent his/her real abilities due to test complexity, environment or student mental state when the test was taken, etc. Additionally, item-by-item reporting disregards the variations in proficiency of the subgroups to which the set of items was administered. The use of item response theory scaling helps get around some of the drawbacks of correct number responses approach. When a certain ability is required to reply to a group of items, the response patterns should display regularities that can be modelled utilizing the underlying similarities between the items. This enables the estimation of the relationships between proficiency and background characteristics and the description of performance distributions within a community or subpopulation [21]. Therefore, in international studies PVs are used to improve measurement accuracy.

The methodology of the PVs in PISA 2015 survey consists of mathematical calculation of posterior distributions around the reported values and obtaining 10 plausible values from these distributions for every student. Consequently, plausible values can be described as random values drawn from the posterior distributions [16]. Plausible values are included in initial PISA 2015 dataset. In multilevel modelling the best course of action would be to use all of each student's PVs, however, due to software limitations, the usual practise of PV in modelling includes only one of these values. This works' modelling is based on J. Mang et al. [10] research, hence, science literacy performance (dependent variable) is defined as the first PV.

4.3.2 Regressors

Regressors used in this study are a variety of factors from a students' socio-economic background that are considered to influence the outcome of student performance, in this case, the scientific literacy. Considering that PISA data is highly hierarchical, the regressors are presented as student or school level variables. The variables that are believed to influence the performance on a student level are:

- Index of economic, social and cultural status (ESCS) – a derived composite interval scale variable calculated by the indicators parental education, highest parental occupation and home possessions, where the higher index score means the higher student status;
- Whole days of school missed in the last two full weeks (ST062Q01TA) – a variable ranging values from 1 to 4, where: 1 – none, 2 – one or two times, 3 – three or four times, 4 – five or more times;
- Environmental awareness (ENVAWARE) – a derived interval scale variable representing how well student is aware of environmental matters;
- Enjoyment of science (JOYSCIE) – a derived interval scale variable showing students excitement related to science;
- Enjoyment of cooperation (COOPERATE) – a derived interval scale variable measuring student willingness to collaborate;
- Use of ICT at school in general (USESCH) – a derived interval scale variable explaining how often digital devices are used at school in general (for chatting, browsing, posting work online, doing schoolwork and etc.).

Regressors that are believed to influence the performance on a school level are as follows:

- Description of the community in which school is located (SC001Q01TA) – a variable with values: 1 – a village, hamlet or rural area (fewer than 3000 people), 2 – a small town (3000 to about 15 000 people), 3 – a town (15 000 to about 100 000 people), 4 – a city (100 000 to about 1 000 000 people), 5 – a large city (with over 1 000 000 people);
- Index proportion of all teachers ISCED level 6 (PROAT6). The proportion was transformed into percentage in this research;
- School Size (SCHSIZE) – total enrolment at school;
- Estimated percentage of students with special needs from national modal grade for 15-year-olds (SC048Q02NA);
- Estimated percentage of students from socio-economic disadvantaged homes from national modal grade for 15-year-olds (SC048Q03NA).

Uppercase items in the parentheses denote the variable coding in PISA databases. It is worthy to mention that, for the whole data, means and standard deviations for derived variables are respectably 0 and 1, but for Lithuanian data these parameters are slightly different, as the derived variables were standardised for all OECD countries and partner counties/economies. When speaking about derived variable (ENVAWARE, JOYSCIE, COOPERATE, USESCH) scores, OECD emphasizes that 'It is possible to interpret these scores by comparing individual scores or group mean scores to the OECD mean, but the individual scores do not reveal anything about the actual item responses and it is impossible to determine from scale score values to what extent respondents endorsed the items used for the measurement of the latent construct. Negative values on the index do not imply that students responded negatively to the underlying question. Rather, students with negative scores are those who responded less positively than the average student across OECD countries. Likewise, students with positive scores are those who responded more positively than the average student in OECD countries.' [19]. The summary of Lithuanian derived variables is presented in Table 1.

Derived variables	Min	Max	Mean	St. deviation
ESCS	-4.05	3.38	-0.07	0.87
ENVAWARE	-3.38	3.28	0.42	1.35
JOYSCIE	-2.11	2.16	0.30	1.15
COOPERATE	-3.33	2.29	0.09	1.16
USESCH	-1.67	3.63	0	1.09

Table 1: Summary of Lithuanian data derived variables

For more information about regressors see [19].

4.4 Weights

Although the students participating in the final PISA sample for a specific nation or economy were selected at random, there are different student and school selection probabilities. To make sure that each participating student accurately represents the right number of students throughout the whole PISA sample, survey weights must be applied into the analysis. The relative contribution of each participating unit to the final population estimate is managed by sampling weights [22]. The weight W_{ji} for student i in school j is made up of two base weights, five adjustment factors, the school base weight and the within-school base weight and can be denoted as:

$$W_{ji} = t_{2ji} f_{1j} f_{2ji} f_{1ji}^{A} t_{1j} w_{2ji} w_{1j} \tag{2}$$

Where:

- w_{1j} (the school base weight) is provided as the reciprocal of the likelihood that school j will be included in the sample;
- w_{2ji} (the within-school base weight) is stated as the reciprocal of the likelihood that student i will be chosen from the chosen school j;
- f_{1j} is an adjusting factor to make up for the absence of other schools that are relatively similar to school j in character (not already made up for by the involvement of replacement schools);
- f_{1ji}^{A} is a correction factor designed to make up for schools in some participating nations where the evaluation only included 15-year-old students who were enrolled in the modal grade for 15-year-old children;
- f_{2ji} is a correction factor used to account for children who did not participate but who were in the same explicit strata and school non-response cell, as well as the same high/low grade and gender categories, depending on the sample size;

- t_{1j} is a trimming factor for the school base weight that is used to lower w_{1j} values that are surprisingly big;
- t_{2ji} is a final student weight trimming factor that is used to lower the weights of students whose weight components have unusually high values.

Student weights W_{ji} are generally applied to student level data for analysis and in this work it is used for Exploratory analysis, multilevel analysis cannot be done using weight W_{ji} because the first stage's condition of equal probability sampling is violated in PISA studies. There are many different approaches to solving this issue; nevertheless, they are all essentially based on a design-based estimating model that uses the pseudo maximum likelihood technique and modified weights for the respective hierarchies. There are a number of various methods for applying and scaling sample weights in hierarchical models that are advocated, but no study before the research of J. Mang et al. [10] has compared them to show which works the best and should thus be chosen. J. Mang et al. [10] used PISA 2015 data and revealed that one of the better techniques for practical implications involves utilizing only school weights w_{1j} and, due to its ease of use, it is also the most favorable one. Therefore, school weights w_{1j}, reported by PISA, are used in this study's multilevel modelling as level two weights, while no weights are used at student level. For further information about survey weighting methodology see [22].

4.5 Normality assumption

Typically, data used in HLM is in interval scales. Theoretically, the error terms for the interval variables are assumed to be normally distributed. However, this presumption is typically overlooked in research (like it is seen in [10]) and is seldom satisfied in practice. Although different transformations of the variables (Z-Score Standardization and Log Transformation) were used in our investigation to try to guarantee that the residuals were normal, the histograms did not closely resemble a normal distribution and Shapiro-Wilk test showed that the null hypothesis of normal distribution should be rejected. Figure 2 below illustrates the original distribution and transformed distributions of variable use of ICT at school in general (USESCH) for 9th graders.

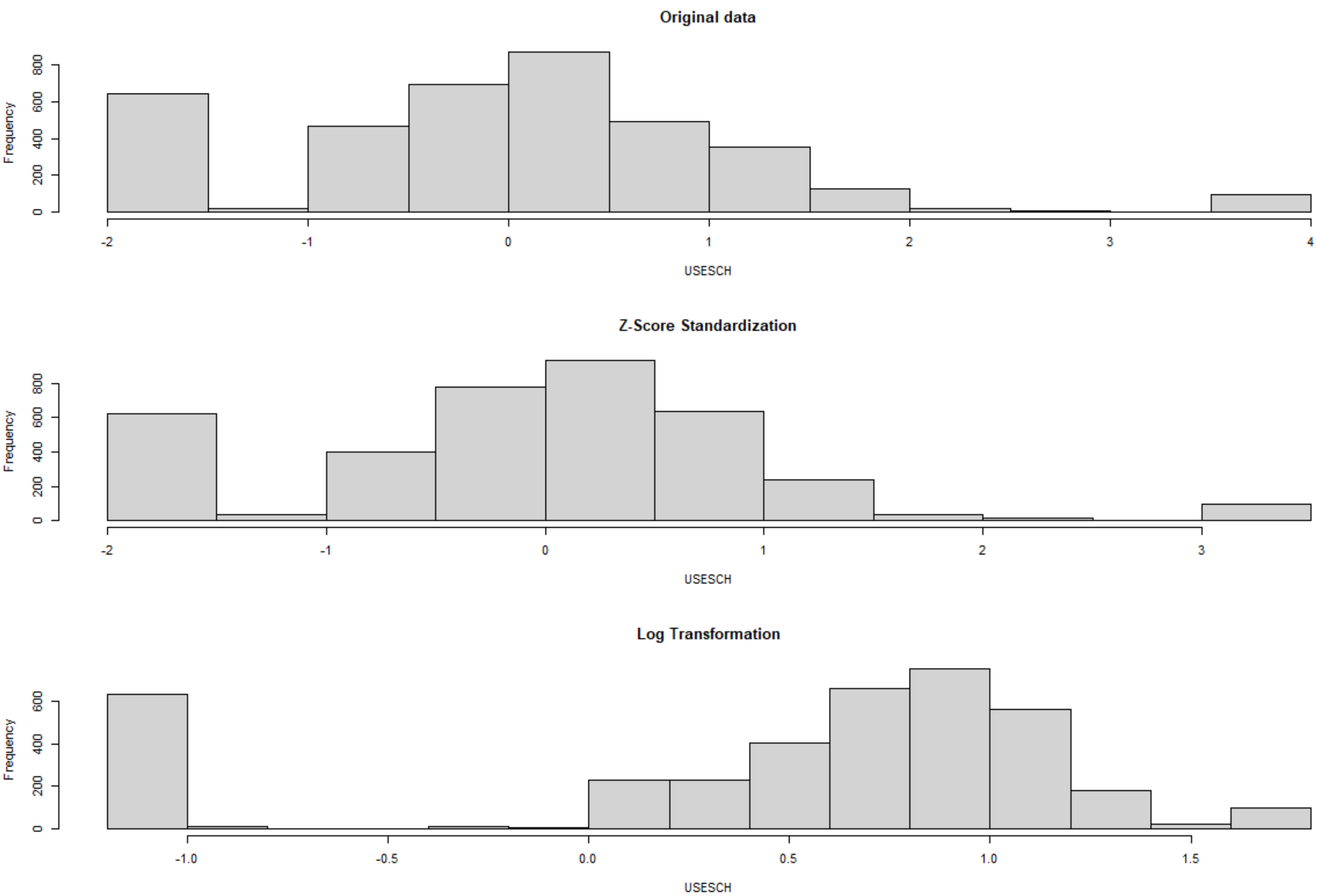

Figure 2: Histograms of variable use of ICT at school in general

5 Real data analysis

5.1 Exploratory analysis

In Lithuania, a sample of 6525 students represented a total of 32 097 15-year-old student population enrolled to grade 7 or above. Based on PISA 2015 data, the weighted average score for scientific literacy in Lithuania was 475.41 with female students slightly outperforming males. Table 5 below represents the number, weighted number and weighted mean score statistics of the sampled students within different levels of analysis.

In this analysis students are defined as socio-economically advantaged (disadvantaged) if they are amongst the 25% of students with the highest (lowest) values in the ESCS index in their country and socio-economically average if they belong to the middle 50%. This definition is used by OECD [25] and the data suggests that this index, calculated by PISA, has an enormous impact on scientific performance, while advantaged students scored almost 48 and 82 points more than average and disadvantaged kids, respectively.

Another level of analysis is the size of a settlement where the school is located. The city is defined as having a total population of approximately 100 000 to 1 000 000 people, town – 15 000 to 100 000 residents, small town – 3000 to 15 000 residents and village less than 3000 people. Students from cities scored well above average, while students from villages received much worse results than their peers from bigger settlements. The significant difference between students from cities and villages approves the idea of A. Zabulionis that 'the achievement gap between urban and rural schools is widening' [32]. Generally, the bigger the settlement, the higher the results were on average. This may be partially explained by the varying proportions of socio-economical statuses within different size settlements, presented in Table 2.

Settlement size	ESCS adv.	ESCS avg.	ESCS disadv.
City	36.2%	49.6%	14.2%
Town	21.9%	57.3%	20.8%
Small town	21.1%	50.4%	28.5%
Village	9.8%	44.4%	45.8%

Table 2: Proportions of students' socio-economical statuses

Most students, to be exact 83.11%, taking PISA test were 9th graders with an average score almost identical to the country level score. Second by size stratum was 10th graders, who represented 14.03% of the sample with an average score of 497.43. Moreover, there is a trend towards significantly higher scores in the higher grades. To investigate the further differences between the two major strata of grades 9 and 10, two different models are presented in section 5.2.

According to a more thorough examination of school type levels, students from the gymnasiums received the highest scores while students from vocational schools scored the least. In addition, secondary schools outperformed basic schools, probably because basic schools only go up to grade 10 and their students, presumably, are less motivated to achieve academically. And although gymnasiums outperformed progymnasiums by a wide margin, this may be due to the fact that the majority of students in progymnasiums were eighth graders and in gymnasiums ninth graders. A more detailed Table 3 illustrates the distribution of grades in different types of schools.

School type	7 grade	8 grade	9 grade	10 grade	11 grade
Gymnasium	0%	0%	86.8%	13.1%	0.01%
Progymnasium	2.3%	97.7%	0%	0%	0%
Secondary	0.1%	3.3%	76.1%	20.4%	0.1%
Basic	0.4%	5.6%	81.1%	12.9%	0%
Vocational	0%	0%	87.3%	12.7%	0%

Table 3: Proportions of grades

Intuitively, schools that have admissions (record of performance or placement tests) attract better students. Therefore, it is not a coincidence that schools with special admissions outperformed other schools. In Lithuania, admissions are often to grade 9, so is it possible that these schools had higher scores because they only have grade 9 or above, or simply because they have more students from higher grades? The following Table 4 denies this idea because schools with different practices have very similar grade proportions.

Admission practise	7 grade	8 grade	9 grade	10 grade	11 grade
Special admission: Never	0.1%	3.6%	81.2%	15.0%	0.1%
Special admission: Sometimes	0.1%	2.1%	83.8%	13.9%	0.1%
Special admission: Always	0.1%	1.8%	85.5%	12.5%	0.1%

Table 4: Proportions of grades

Level	Number	Weighted number	Mean
Lithuania	6525	29914.59	475.41
Males	3324	15181.5	471.77
Females	3201	14733.1	479.16
ESCS advantaged	1583	7376.42	520.09
ESCS average	3167	14462.67	472.37
ESCS disadvantaged	1584	7198.10	438.36
City	2670	11328.63	499.31
Town	1112	5912.85	478.09
Small town	1372	6394.92	461.72
Village	1371	6278.20	443.70
Modal grade +2	5	6.92	543.26
Modal grade +1	916	3280.52	497.43
Modal grade	5423	25813.63	475.21
Modal grade -1	174	786.93	393.84
Modal grade -2	7	26.59	349.91
Gymnasium	4343	21455.81	492.40
Progymnasium	86	384.11	410.97
Secondary	1010	2743.25	468.56
Basic	1015	4887.36	421.19
Vocational	71	444.07	349.46
Special admission: Never	2862	11942.01	451.03
Special admission: Sometimes	2065	9756.25	469.32
Special admission: Always	1566	8058.55	518.69

Table 5: Summary of Exploratory data

5.2 Modelling

In this research RStudio software with R version 4.2.1 is employed. As it was mentioned before, this software has some limitations. To authors knowledge, R does not have a package to perform HLM with all plausible values together with multilevel weights, therefore, there was a choice of performing student level modelling with all plausible values and student level weights or conducting multilevel modelling with 1 PV and multilevel weights. The latter option is used to perform HLM with a state-of-the-art WeMix package (publicized on 2022-10-05) because this package is unique in employing techniques for hierarchical linear models that use unequal weights at various levels. For linear models, the model is evaluated with a weighted version of the estimating equations used in lme4 package, a package that provides functions for fitting and analysing mixed models. The practical use of this package is followed by a recommendation from an article by J. Mang et al. [10] to use only school level weights combined with dummy student level weights.

To check if hierarchical structure should be taken into account, the unconditional model is run first, as the schools are suspected of having a significant impact on ninth and tenth graders achievement (first PV of science literacy). It is also worth noting that the initial number of observations from different classes is reduced as the missing values of the regressors are cleaned. The results reveal that higher grade strata have higher variances and the intercept has a higher value, as this should be suspected. Moreover, both intraclass correlation coefficients confirm that hierarchical structure should be accounted. The summary of separate unconditional models is presented in Table 6.

Models	N	School level variance	Student level variance	ICC	Intercept
9th graders	3797	1194	5714	17.28%	459.15
10th graders	639	1373	6232	18.06%	481.97

Table 6: Null models for 9th and 10th graders

The following part examines the multilevel model. To begin with, the correlation matrices of regressors are calculated for separate grades and are presented in Tables 7 and 8. Both 9 and 10 grade initial models have the following form:

$$\begin{aligned}
\text{Student level: } Y_{ij} &= \beta_{0j} + \beta_{1j}ESCS_{ij} + \beta_{2j}ST062Q01TA_{ij} + \beta_{3j}ENVAWARE_{ij} + \\
&+ \beta_{4j}JOYSCIE_{ij} + \beta_{5j}COOPERATE_{ij} + \beta_{6j}USESCH_{ij} + e_{ij}; \\
\text{School level: } \beta_{0j} &= \gamma_{00} + \gamma_{01}SC001Q01TA_{1j} + \gamma_{02}PROAT6_{2j} + \gamma_{03}SCHSIZE_{3j} + \\
&+ \gamma_{04}SC048Q02NA_{4j} + \gamma_{05}SC048Q03NA_{5j} + u_{0j};; \\
\beta_{1j} &= \gamma_{10} + u_{1j}; \\
\beta_{2j} &= \gamma_{20} + u_{2j}; \\
\beta_{3j} &= \gamma_{30} + u_{3j}; \\
\beta_{4j} &= \gamma_{40} + u_{4j}; \\
\beta_{5j} &= \gamma_{50} + u_{5j}; \\
\beta_{6j} &= \gamma_{60} + u_{6j}.
\end{aligned}$$

Therefore, the pooled model is:

$$\begin{aligned}
Y_{ij} &= \gamma_{00} + \gamma_{01}SC001Q01TA_{1j} + \gamma_{02}PROAT6_{2j} + \gamma_{03}SCHSIZE_{3j} + \gamma_{04}SC048Q02NA_{4j} + \\
&+ \gamma_{05}SC048Q03NA_{5j} + \gamma_{10}ESCS_{ij} + \gamma_{20}ST062Q01TA_{ij} + \gamma_{30}ENVAWARE_{ij} + \\
&+ \gamma_{40}JOYSCIE_{ij} + \gamma_{50}COOPERATE_{ij} + \gamma_{60}USESCH_{ij} + e_{ij} + u_{0j} + \\
&+ ESCS_{ij}u_{1j} + ST062Q01TA_{ij}u_{2j} + ENVAWARE_{ij}u_{3j} + JOYSCIE_{ij}u_{4j} + \\
&+ COOPERATE_{ij}u_{5j} + USESCH_{ij}u_{6j}.
\end{aligned}$$

	ESCS	ST062Q01TA	ENVAWARE	JOYSCIE	COOPERATE	USESCH	SC001Q01TA	PROAT6	SCHSIZE	SC048Q02NA	SC048Q03NA
ESCS	1.00	-0.08	0.20	0.10	0.13	0.04	0.32	0.15	0.33	-0.23	-0.27
ST062Q01TA	-0.08	1.00	-0.11	-0.12	-0.10	0.10	-0.02	-0.02	-0.06	0.07	0.06
ENVAWARE	0.20	-0.11	1.00	0.33	0.24	0.01	0.05	0.03	0.10	-0.08	-0.10
JOYSCIE	0.10	-0.12	0.33	1.00	0.27	0.03	-0.03	0.03	0.03	-0.02	-0.05
COOPERATE	0.13	-0.10	0.24	0.27	1.00	0.00	0.03	0.02	0.09	-0.06	-0.07
USESCH	0.04	0.10	0.01	0.03	0.00	1.00	-0.06	-0.06	-0.06	0.04	0.04
SC001Q01TA	0.32	-0.02	0.05	-0.03	0.03	-0.06	1.00	0.25	0.53	-0.34	-0.48
PROAT6	0.15	-0.02	0.03	0.03	0.02	-0.06	0.25	1.00	0.11	-0.04	-0.15
SCHSIZE	0.33	-0.06	0.10	0.03	0.09	-0.06	0.53	0.11	1.00	-0.40	-0.44
SC048Q02NA	-0.23	0.07	-0.08	-0.02	-0.06	0.04	-0.34	-0.04	-0.40	1.00	0.55
SC048Q03NA	-0.27	0.06	-0.10	-0.05	-0.07	0.04	-0.48	-0.15	-0.44	0.55	1.00

Table 7: Correlation matrix for grade 9

	ESCS	ST062Q01TA	ENVAWARE	JOYSCIE	COOPERATE	USESCH	SC001Q01TA	PROAT6	SCHSIZE	SC048Q02NA	SC048Q03NA
ESCS	1.00	-0.13	0.20	0.07	0.11	-0.04	0.40	0.11	0.35	-0.20	-0.31
ST062Q01TA	-0.13	1.00	-0.12	-0.12	-0.13	0.14	-0.14	-0.07	-0.08	0.04	0.04
ENVAWARE	0.20	-0.12	1.00	0.37	0.28	-0.01	0.07	0.00	0.12	0.03	-0.02
JOYSCIE	0.07	-0.12	0.37	1.00	0.28	-0.01	-0.03	0.03	0.03	-0.01	0.06
COOPERATE	0.11	-0.13	0.28	0.28	1.00	0.06	0.10	0.03	0.09	-0.05	-0.04
USESCH	-0.04	0.14	-0.01	-0.01	0.06	1.00	-0.09	-0.10	-0.08	0.07	0.08
SC001Q01TA	0.40	-0.14	0.07	-0.03	0.10	-0.09	1.00	0.22	0.51	-0.37	-0.48
PROAT6	0.11	-0.07	0.00	0.03	0.03	-0.10	0.22	1.00	0.11	-0.03	-0.19
SCHSIZE	0.35	-0.08	0.12	0.03	0.09	-0.08	0.51	0.11	1.00	-0.33	-0.33
SC048Q02NA	-0.20	0.04	0.03	-0.01	-0.05	0.07	-0.37	-0.03	-0.33	1.00	0.55
SC048Q03NA	-0.30	0.04	-0.02	0.06	-0.04	0.08	-0.48	-0.19	-0.33	0.55	1.00

Table 8: Correlation matrix for grade 10

The final models for ninth and tenth graders are constructed by removing one insignificant value with the largest p-value at the time until only the significant regressors are left and taking into account the correlation matrix. The final models and their logic are presented in the following Results section.

6 Results

To determine the multilevel models' effectiveness conditional change of first level variance (CCV) is calculated, that is:

$$CCV = \frac{\sigma^2_{null} - \sigma^2_{new}}{\sigma^2_{null}}, \tag{3}$$

where σ^2_{null} is the student level variance estimate for null model and σ^2_{new} is the student level variance estimate for the new model. CCV is used to determine how much the differences are reduced compared to the unconditional model. In our models CCV is quite large (18.15%-21.95%), which demonstrates that compared to null model, new models are much better.

Both initial models in Table 9 consist of 6 student level and 5 school level variables. The final models are obtained by removing one insignificant variable with the highest p-value per iteration until only the significant regressors are left. It is important to note that variable that describes the community in which school is located (SC001Q01TA) has a negative value in grade 9 model. According to the Table 5 this does not make sense, since the bigger the settlement the school is located in, the better the results, on average. This variable is correlated with other variables that are associated with the size of school (schools in larger settlements are often bigger) and that might cause its odd value. Also, variable est. percentage of students with special needs (SC048Q02NA) has an odd value in both initial models, but that might be due to its correlation with variable estimated percentage of students from socio-economic disadvantaged homes (SC048Q03NA).

Effects & Statistics	Grade 9 model	Grade 10 model
Intercept	476.93*** (8.84)	465.26*** (12.66)
Index of economic, social and cultural status (ESCS)	14.21*** (1.57)	26.60*** (3.27)
Whole days of school missed in the last two full weeks (ST062Q01TA)	-15.72*** (1.79)	-20.64*** (4.00)
Environmental awareness (ENVAWARE)	14.18*** (0.97)	11.16*** (2.39)
Enjoyment of science (JOYSCIE)	6.68*** (1.06)	6.15. (3.21)
Enjoyment of cooperation (COOPERATE)	4.96*** (1.14)	6.02* (3.07)
Use of ICT at school in general (USESCH)	-13.32*** (1.16)	-6.81* (2.91)
Description of the community in which school is located (SC001Q01TA)	-4.03. (2.06)	9.78** (3.01)
Index proportion of all teachers ISCED level 6 (PROAT6)	6.82** (2.42)	7.91. (4.10)
School Size (SCHSIZE)	0.05*** (0.01)	0.03* (0.01)
Est. percentage of students with special needs (SC048Q02NA)	0.13 (0.26)	0.77 (0.48)
Est. percentage of students from socio-economic disadvantaged homes (SC048Q03NA)	-0.39** (0.14)	-0.19 (0.21)
School level variance	366.6	174.5
Student level variance	4665.7	4864.1
CCV	18.35%	21.95%

Note: Standard errors are in parentheses. Predictors abbreviations are in square brackets.
Significance codes: .p < .1, *p < .05, **p < .01, ***p < .001.

Table 9: Initial models

The final models presented in Table 10 show that there are regressors that are statistically significant at both individual and school level. Although, there are more statistically significant variables in grade 9 model, CCV metric indicates that grade 10 model 'captures' more variance. To be exact, in comparison to zero models, the differences in student results unexplained by the final models are reduced by 18.15% and 20.81%, respectively. Also, in both final models the school level variance is significantly reduced.

As it was expected, the index of economic, social and cultural status (ESCS) and whole days of school missed in the last two full weeks (ST062Q01TA) have corresponding positive and negative signs and contribute significantly to the scientific literacy. Furthermore, the more a student is aware of environmental matters (ENVAWARE), the better his/her results are on average. Also, the enjoyment of activities related to science and cooperation (JOYSCIE and COOPERATE) is important, since students who are experiencing positive emotions are more likely to achieve better results. However, it is interesting that the variable corresponding to the enjoyment of science is not significant in grade 10 model. One of the unexpected results is the negative coefficient of the variable corresponding to the use of ICT at school in general (USESCH) in both models. It can be assumed that students are not using digital devices properly (browse social networks, play games and probably do not use them for learning purposes).

One school level regressor regarding the estimated percentage of students with special needs from national modal grade for 15-year-olds (SC048Q02NA) is removed from both final models because the variable appears to not be significant. Therefore, four variables are left of which index proportion of all teachers ISCED level 6 (PROAT6) is significant for both models, est. percentage of students from socio-economic disadvantaged homes from national modal grade for 15-year-olds (SC048Q03NA) and school size (SCHSIZE) are significant only for 9th graders and the variable corresponding to the description of the community in which school is located (SC001Q01TA) is significant only in grade 10 model. Possible assumption can be made about variables corresponding to the size of school and description of the settlement size where the school is located. First one is only significant in grade 9 model, the other is significant in grade 10 model. One can presume that these regressors are very similar because they strongly correlate (cities have bigger schools than villages on average and so on).

Effects & Statistics	Grade 9 model	Grade 10 model
Intercept	469.44*** (7.63)	473.99*** (9.89)
Index of economic, social and cultural status (ESCS)	13.77*** (1.54)	27.81*** (3.19)
Whole days of school missed in the last two full weeks (ST062Q01TA)	-15.81*** (1.78)	-20.95*** (4.02)
Environmental awareness (ENVAWARE)	14.24*** (0.97)	13.29*** (2.31)
Enjoyment of science (JOYSCIE)	6.83*** (1.07)	NA
Enjoyment of cooperation (COOPERATE)	5.02*** (1.13)	7.07* (2.90)
Use of ICT at school in general (USESCH)	-13.27*** (1.16)	-6.99* (3.00)
Description of the community in which school is located (SC001Q01TA)	NA	11.43*** (2.68)
Index proportion of all teachers ISCED level 6 (PROAT6)	6.09** (2.34)	8.85* (3.96)
School Size (SCHSIZE)	0.04*** (0.01)	NA
Est. percentage of students from socio-economic disadvantaged homes (SC048Q03NA)	-0.30* (0.13)	NA
School level variance	361.9	183.6
Student level variance	4677.1	4935.4
CCV	18.15%	20.81%

Note: Standard errors are in parentheses. Predictors abbreviations are in square brackets.
Significance codes: .p < .1, *p < .05, **p < .01, ***p < .001.
NA value means the regressor is not used in the final model.

Table 10: Final models

Estimated grade 9 model is:

$$\begin{aligned} Y_{ij} &= 469.44 + 13.77ESCS - 15.81ST062Q01TA + 14.24ENVAWARE+ \\ &+ 6.83JOYSCIE + 5.02COOPERATE - 13.27USESCH+ \\ &+ 6.09PROAT6 + 0.04SCHSIZE - 0.30SC048Q03NA; \end{aligned} \tag{4}$$

and estimated grade 10 model is:

$$\begin{aligned} Y_{ij} &= 473.99 + 27.81ESCS - 20.95ST062Q01TA + 13.29ENVAWARE+ \\ &+ 7.07COOPERATE - 6.99USESCH + 11.43SC001Q01TA + 8.85PROAT6. \end{aligned} \tag{5}$$

As it is seen from Equation 4 for 9th graders, 1 unit increase in index of economic, social and cultural status increases scientific literacy score by 13.77 points. Other variables that also look influential are environmental awareness and use of ICT at school. Their 1 unit increase changes average literacy score by 14.24 and -13.27 points accordingly. Another very important variable seems to be the school size. Since the average Lithuanian school in PISA 2015 dataset has 328 students, this would increase the average score by 13.12 points. Other variable that seems to make quite much difference is est. percentage of students from socio-economic disadvantaged homes, since the average percentage of 15-year-olds with socio-economic disadvantage in the Lithuanian modal grade is 34.24%, the average score would decrease by 10.27 points. However, variable index proportion of all teachers ISCED level 6 does not seem to make much impact as, for instance, a school with 0.29% (mean for Lithuanian schools in PISA 2015 dataset) of teachers with ISCED level 6 education would only increase the average scientific literacy score by only 1.77 points.

Equation 5 for 10th graders shows even greater impact of index of economic, social and cultural status, as 1 unit increase in the index increases scientific literacy score by 27.81 points. Another variable that looks to be influential is environmental awareness. It has similar coefficient as in the model for 9th graders and its 1 unit increase changes literacy score by 13.29 points. Even though grade 10 final model does not have variable school size, it has a regressor that describes the settlement size where the school is located (SC001Q01TA) and these variables are related. The variable describing the settlement size where the school is located seems to be very impactful because, according to the equation, living in a city (100 000 to about 1 000 000 people) increases the average score by 45.72 points while living in a village (fewer than 3000 people) – 11.43 points. Once again, in grade 10 model index proportion of all teachers ISCED level 6 does not seem to be very influential since a school with 0.29% of teachers with this level of education would increase the scientific literacy score by 2.57 points.

7 Conclusions

This study is carried out to identify the links/impacts of socio-economic background on students' scientific literacy using PISA 2015 data and performing hierarchical linear modelling with R software. Two different models are created for grades 9 and 10 in order to understand which variables can impact the academic performance on different levels. The work is one of the first that investigates socio-economic background and academical performance relationship in Lithuanian educational system through data modelling, as previously only descriptive and exploratory data analysis has been carried out on this topic. Furthermore, the research includes the use of the newest study's findings about the practical implications of sampling weights in multilevel modelling, the plausible value theory and the state-of-the-art package WeMix used for mixed-effects models that includes weights at every level. One of the HLM assumptions is a normality of residuals. This assumption is rarely met in reality and usually ignored in research. In our study, attempts were made to ensure normality of the residuals by transforming the variables in various ways (Z-Score Standardization and Log Transformation) but the histograms were not close to a normal distribution (this could be due to the large number of observations). Therefore, our results should be viewed with a certain amount of caution but nevertheless might be used as basis for further research.

The results show that there are regressors that are significant for student literacy. The significant regressors obtained are the index of economic, social and cultural status, days of school missed, environmental awareness, enjoyment of science, enjoyment of cooperation, use of ICT at school in general, settlement size in which school is located, the index proportion of all teachers ISCED level 6, school size and estimated percentage of students from socio-economic disadvantaged homes. Although variables vary in grade 9 and grade 10 models, these models are essentially quite similar. Furthermore, it can be noted from Equations 4 and 5 that variables index of economic, social and cultural status, environmental awareness are very influential in both grade 9 and 10 models. Moreover, variable school size is only significant in grade 9 model, while the regressor settlement size where the school is located is significant only in grade 10 model but these regressors are highly related and both make a great impact on student achievement. The variable that does not seem to be crucial in both models is index proportion of all teachers ISCED level 6 as for a school with 0.29% (mean for Lithuanian schools in PISA 2015 dataset) of teachers with ISCED level 6 education it would only increase the average scientific literacy score by 1.77-2.57 points, depending on the model. It can also be seen that one of the regressors (index of economic, social and cultural status) is directly related to socio-economic perspective, others such as days of school missed, use of ICT at school in general, settlement size in which school is located, index proportion of all teachers ISCED level 6, school size and est. percentage of students from socio-economic disadvantaged homes are related partially. The remaining variables like environmental awareness, enjoyment of science and enjoyment of cooperation are more general and are connected on a pedagogical level. Clearly, these factors are important if we want to compensate for the educational achievements of socio-economically disadvantaged students, as we cannot change a student's SES or the place in which he or she lives, but we can modify other regressors that are important to compensate for differences in the achievements of students of different SES.

8 Appendix

```
library(memisc)
library("readr")
library("data.table")
library(haven)
library(dplyr)
library(magrittr)
library(foreign)
library(knitr)
library(afex)
library(nlme)
library(msm)
library(car)
library(AICcmodavg)
library(Hmisc)
library(expss)
library(psych)
library(tidyr)
library(mitools)
library(mitml)
library(corrplot)
library(WeMix)
library(plotrix)


path = file.path("C:/Users/pauli/OneDrive/Stalinis_kompiuteris/Master/CY6_MS_CMB_STU_QQQ.sav")
dataset_student = read_sav(path)
dataset_student_filtered<-filter(dataset_student, CNT =="LTU")


path2 = file.path("C:/Users/pauli/OneDrive/Stalinis_kompiuteris/Master/CY6_MS_CMB_SCH_QQQ.sav")
dataset_school = read_sav(path2)
dataset_school_filtered<-filter(dataset_school, CNT =="LTU")


dataset_filtered_merged <- merge(dataset_student_filtered, dataset_school_filtered, by="CNTSCHID")




##########
#Means, std dev. of derived variables
##########
summary(dataset_student_filtered$ESCS)
summary(dataset_student_filtered$ENVAWARE)
summary(dataset_student_filtered$JOYSCIE)
summary(dataset_student_filtered$COOPERATE)
summary(dataset_student_filtered$USESCH)

sd(dataset_student_filtered$ESCS, na.rm = TRUE)
sd(dataset_student_filtered$ENVAWARE, na.rm = TRUE)
sd(dataset_student_filtered$JOYSCIE, na.rm = TRUE)
sd(dataset_student_filtered$COOPERATE, na.rm = TRUE)
sd(dataset_student_filtered$USESCH, na.rm = TRUE)
```

```
##############
#EDA
##############

dataset_filtered_merged$SCIE_REZ_weights<-((dataset_filtered_merged$PV1SCIE*dataset_filtered_merged
    $W_FSTUWT)+
                                    (dataset_filtered_merged$PV2SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV3SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV4SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV5SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV6SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV7SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV8SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV9SCIE*dataset_filtered_merged$W_
                                        FSTUWT)+
                                    (dataset_filtered_merged$PV10SCIE*dataset_filtered_merged$W_
                                        FSTUWT))/10

sum(dataset_filtered_merged$SCIE_REZ_weights)/sum(dataset_filtered_merged$W_FSTUWT)


#Gender
rez_sum_lytis<-aggregate(x = dataset_filtered_merged$SCIE_REZ_weights, # Specify data column
                    by = list(dataset_filtered_merged$ST004D01T), # Specify group indicator
                    FUN = sum) # Specify function (i.e. mean)
wt_sum_lytis<-aggregate(x = dataset_filtered_merged$W_FSTUWT, # Specify data column
                    by = list(dataset_filtered_merged$ST004D01T), # Specify group indicator
                    FUN = sum)

rez_sum_lytis[,1]
rez_sum_lytis[,2]/wt_sum_lytis[,2]
table(dataset_filtered_merged$ST004D01T) #1-Female


#ESCS
dataset_filtered_merged$ESCS_class <- as.numeric(cut2(dataset_filtered_merged$ESCS, g=4))
dataset_filtered_merged$ESCS_class[dataset_filtered_merged$ESCS_class=="3"]<-"2"
dataset_filtered_merged$ESCS_class[dataset_filtered_merged$ESCS_class=="4"]<-"3"

rez_sum_ESCS<-aggregate(x = dataset_filtered_merged$SCIE_REZ_weights, # Specify data column
                    by = list(dataset_filtered_merged$ESCS_class), # Specify group indicator
                    FUN = sum) # Specify function (i.e. mean)
wt_sum_ESCS<-aggregate(x = dataset_filtered_merged$W_FSTUWT, # Specify data column
                    by = list(dataset_filtered_merged$ESCS_class), # Specify group indicator
                    FUN = sum)
rez_sum_ESCS[,2]/wt_sum_ESCS[,2]
table(dataset_filtered_merged$ESCS_class)

#Settlement size
rez_sum_miestas<-aggregate(x = dataset_filtered_merged$SCIE_REZ_weights, # Specify data column
                    by = list(dataset_filtered_merged$SC001Q01TA), # Specify group indicator
```

```
                              FUN = sum) # Specify function (i.e. mean)
wt_sum_miestas<-aggregate(x = dataset_filtered_merged$W_FSTUWT, # Specify data column
                             by = list(dataset_filtered_merged$SC001Q01TA), # Specify group indicator
                             FUN = sum)
rez_sum_miestas[,2]/wt_sum_miestas[,2]
table(dataset_filtered_merged$SC001Q01TA)


#ESCS distribution by city size

df_proporcijos<-dataset_filtered_merged
df_proporcijos<-df_proporcijos[!is.na(df_proporcijos$ESCS_class),]


df_didmiestis<-df_proporcijos
df_didmiestis <-filter(df_didmiestis,SC001Q01TA==4)
df_didmiestis %>% group_by(df_didmiestis$ESCS_class) %>%summarise(Percentage=n()/nrow(.))

df_miestas<-df_proporcijos
df_miestas <-filter(df_miestas,SC001Q01TA==3)
df_miestas %>% group_by(df_miestas$ESCS_class) %>%summarise(Percentage=n()/nrow(.))

df_miestelis<-df_proporcijos
df_miestelis <-filter(df_miestelis,SC001Q01TA==2)
df_miestelis %>% group_by(df_miestelis$ESCS_class) %>%summarise(Percentage=n()/nrow(.))

df_kaimas<-df_proporcijos
df_kaimas <-filter(df_kaimas,SC001Q01TA==1)
df_kaimas %>% group_by(df_kaimas$ESCS_class) %>%summarise(Percentage=n()/nrow(.))




#Results by grades
rez_sum_klase<-aggregate(x = dataset_filtered_merged$SCIE_REZ_weights, # Specify data column
                            by = list(dataset_filtered_merged$GRADE), # Specify group indicator
                            FUN = sum) # Specify function (i.e. mean)
wt_sum_klase<-aggregate(x = dataset_filtered_merged$W_FSTUWT, # Specify data column
                           by = list(dataset_filtered_merged$GRADE), # Specify group indicator
                           FUN = sum)

rez_sum_klase[,1]
rez_sum_klase[,2]/wt_sum_klase[,2]
table(dataset_filtered_merged$GRADE)




#Results by school type
dataset_filtered_merged$PROGN[dataset_filtered_merged$PROGN=="04400003"]<-"04400002" #sujungiu
    Secondary School (Lower Secondary) ir Secondary School (Upper Secondary)
dataset_filtered_merged$PROGN[dataset_filtered_merged$PROGN=="04400006"]<-"04400005" #sujungiu
    Lower Gymnasium ir Upper Gymnasium
rez_sum_mokykl<-aggregate(x = dataset_filtered_merged$SCIE_REZ_weights, # Specify data column
                             by = list(dataset_filtered_merged$PROGN), # Specify group indicator
                             FUN = sum) # Specify function (i.e. mean)
wt_sum_mokykl<-aggregate(x = dataset_filtered_merged$W_FSTUWT, # Specify data column
```

```
                    by = list(dataset_filtered_merged$PROGN), # Specify group indicator
                    FUN = sum)

rez_sum_mokykl[,2]/wt_sum_mokykl[,2]
table(dataset_filtered_merged$PROGN)


df_BS<-dataset_filtered_merged
df_BS <-filter(df_BS,PROGN=="04400001") #BASIC SCHOOL
df_BS %>% group_by(df_BS$GRADE) %>%summarise(Percentage=n()/nrow(.))

df_SS<-dataset_filtered_merged
df_SS <-filter(df_SS,PROGN=="04400002") #Secondary
df_SS %>% group_by(df_SS$GRADE) %>%summarise(Percentage=n()/nrow(.))

df_PG<-dataset_filtered_merged
df_PG <-filter(df_PG,PROGN=="04400004") #Progymnasium
df_PG %>% group_by(df_PG$GRADE) %>%summarise(Percentage=n()/nrow(.))

df_G<-dataset_filtered_merged
df_G <-filter(df_G,PROGN=="04400005") #Gymnasium
df_G %>% group_by(df_G$GRADE) %>%summarise(Percentage=n()/nrow(.))

df_V<-dataset_filtered_merged
df_V <-filter(df_V,PROGN=="04400007") #Vocational
df_V %>% group_by(df_V$GRADE) %>%summarise(Percentage=n()/nrow(.))


#Student admission to school: Student's record of academic performance (including placement tests)
rez_sum_admission<-aggregate(x = dataset_filtered_merged$SCIE_REZ_weights, # Specify data column
                          by = list(dataset_filtered_merged$SC012Q01TA), # Specify group indicator
                          FUN = sum) # Specify function (i.e. mean)
wt_sum_admission<-aggregate(x = dataset_filtered_merged$W_FSTUWT, # Specify data column
                         by = list(dataset_filtered_merged$SC012Q01TA), # Specify group indicator
                         FUN = sum)

rez_sum_admission[,1]
rez_sum_admission[,2]/wt_sum_admission[,2]
table(dataset_filtered_merged$SC012Q01TA)


df_Never<-dataset_filtered_merged
df_Never <-filter(df_Never,SC012Q01TA==1) #Never
df_Never %>% group_by(df_Never$GRADE) %>%summarise(Percentage=n()/nrow(.))

df_Sometimes<-dataset_filtered_merged
df_Sometimes <-filter(df_Sometimes,SC012Q01TA==2) #Sometimes
df_Sometimes %>% group_by(df_Sometimes$GRADE) %>%summarise(Percentage=n()/nrow(.))

df_Always<-dataset_filtered_merged
df_Always <-filter(df_Always,SC012Q01TA==3) #Always
df_Always %>% group_by(df_Always$GRADE) %>%summarise(Percentage=n()/nrow(.))



##############
#Correlation matrix of all items
```

```
##############
dataset_filtered_merged <- merge(dataset_student_filtered, dataset_school_filtered, by="CNTSCHID")

dataset_filtered_merged <- data.frame(dataset_filtered_merged, stringsAsFactors=FALSE)
dataset_filtered_merged[sapply(dataset_filtered_merged, is.character)] <- lapply(dataset_filtered_merged[
    sapply(dataset_filtered_merged, is.character)], as.factor) #characters to factors
dataset_filtered_merged <- dataset_filtered_merged[,colSums(is.na(dataset_filtered_merged))!=nrow(dataset_
    filtered_merged)] #remove colums with all NA
dataset_filtered_merged<-dataset_filtered_merged[rowSums(is.na(dataset_filtered_merged)) != ncol(dataset_
    filtered_merged), ] #remove rows with all NA


as.numeric.labelled <- function(x) {
  r <- as.numeric(as.factor(x))
  label(r)<-label(x)
  r
}
dataset_filtered_merged <- dataset_filtered_merged %>%
  mutate_if(is.factor,
            .funs = as.numeric.labelled)

df_9 <-filter(dataset_filtered_merged,GRADE==0) #9 grade
df_10 <-filter(dataset_filtered_merged,GRADE==1) #10 grade

res<-cor(names2labels(dataset_filtered_merged$PV1SCIE), names2labels(dataset_filtered_merged), use="
    pairwise.complete.obs")
res<-round(res, 2)
res<-res[,colSums(is.na(res))!=nrow(res)]
res<-data.frame(res)
res <- filter(res, abs(res)>=0.2)
res <- filter(res, res<0.80)



#9graders correlation
res_9<-cor(names2labels(df_9$PV1SCIE), names2labels(df_9), use="pairwise.complete.obs")
res_9<-round(res_9, 2)
res_9<-res_9[,colSums(is.na(res_9))!=nrow(res_9)]
res_9<-data.frame(res_9)
res_9 <- filter(res_9, abs(res_9)>=0.2)
res_9 <- filter(res_9, res_9<0.80)
#10graders correlation
res_10<-cor(names2labels(df_10$PV1SCIE), names2labels(df_10), use="pairwise.complete.obs")
res_10<-round(res_10, 2)
res_10<-res_10[,colSums(is.na(res_10))!=nrow(res_10)]
res_10<-data.frame(res_10)
res_10 <- filter(res_10, abs(res_10)>=0.2)
res_10 <- filter(res_10, res_10<0.80)



######################
#Modelling
#####################
dataset_filtered_merged <- merge(dataset_student_filtered, dataset_school_filtered, by="CNTSCHID")

#Averages for results section
sum(dataset_school_filtered$SCHSIZE*dataset_school_filtered$W_SCHGRNRABWT)/sum(dataset_school_
```

```
    filtered$W_SCHGRNRABWT) #average number of students in school
mean(dataset_school_filtered$SCHSIZE) #not correct

sum(dataset_school_filtered$PROAT6*dataset_school_filtered$W_SCHGRNRABWT)/sum(dataset_school_
    filtered$W_SCHGRNRABWT)*100 #avg % of teachers with level6
mean(dataset_school_filtered$PROAT6) #not correct


dataset_school_filtered_na<-dataset_school_filtered %>% drop_na(SC048Q03NA)
sum(dataset_school_filtered_na$SC048Q03NA*dataset_school_filtered_na$W_SCHGRNRABWT)/sum(
    dataset_school_filtered_na$W_SCHGRNRABWT)
mean(dataset_school_filtered_na$SC048Q03NA) #not correct




dataset_filtered_merged$fake_student_weights<-1 #dummy weights
dataset_filtered_merged$PROAT6<-dataset_filtered_merged$PROAT6*100

df_9 <-filter(dataset_filtered_merged,GRADE==0) #9 grade
df_10 <-filter(dataset_filtered_merged,GRADE==1) #10 grade

df_9<-df_9 %>% drop_na(ESCS, ST062Q01TA, ENVAWARE, JOYSCIE, COOPERATE , USESCH,
    SC001Q01TA, PROAT6, SCHSIZE, SC048Q02NA , SC048Q03NA)
df_10<-df_10 %>% drop_na(ESCS, ST062Q01TA, ENVAWARE, JOYSCIE, COOPERATE , USESCH,
    SC001Q01TA, PROAT6, SCHSIZE, SC048Q02NA, SC048Q03NA)


#################
#Normality
#################
summary(df_9$USESCH)
plot(hist(df_9$USESCH), xlab = "USESCH", ylab = "Frequency", main="")
shapiro.test(df_9$USESCH)

#Z-score
df_9$USESCH_scaled<-scale(df_9$USESCH)
plot(hist(df_9$USESCH_scaled), xlab = "USESCH", ylab = "Frequency", main="")
shapiro.test(df_9$USESCH_scaled)

#Log transformation
df_9$USESCH_log<-log(df_9$USESCH+2)
plot(hist(df_9$USESCH_log), xlab = "USESCH", ylab = "Frequency", main="")
shapiro.test(df_9$USESCH_log)

Zscore<-hist(df_9$USESCH_scaled)
log<-hist(df_9$USESCH_log)
norm<-hist(df_9$USESCH)

par(mfrow=c(3,1))
plot(norm, xlab = "USESCH", ylab = "Frequency", main="Original data")
plot(Zscore, xlab = "USESCH", ylab = "Frequency", main="Z-Score Standardization")
plot(log, xlab = "USESCH", ylab = "Frequency", main="Log Transformation")



#9 grade correlation matrix
```

```
df_9_cor<-data.frame(df_9$ESCS, df_9$ST062Q01TA, df_9$ENVAWARE, df_9$JOYSCIE, df_9$
    COOPERATE,df_9$USESCH,
                df_9$SC001Q01TA,df_9$PROAT6,df_9$SCHSIZE,df_9$SC048Q02NA ,df_9$SC048Q03NA
                )
round(cor(df_9_cor),2)
corrplot(round(cor(df_9_cor, use="pairwise.complete.obs"),2))

#9 grade model
two_level_09_null <- mix(PV1SCIE ~ (1|CNTSCHID), data=df_9, weights=c("fake_student_weights", "W_
    SCHGRNRABWT"))
summary(two_level_09_null) #icc-17,28% #1194 #5714

two_level_09_initial <- mix(PV1SCIE ~ESCS+ ST062Q01TA+ ENVAWARE+ JOYSCIE + COOPERATE +
    USESCH +
              SC001Q01TA + PROAT6+ SCHSIZE+SC048Q02NA+ SC048Q03NA+(1|CNTSCHID), data=df
                _9, weights=c("fake_student_weights", "W_SCHGRNRABWT"))
summary(two_level_09_initial) #366.6 4665.7
(5714-4665.7)/5714 # 18,35%
waldTest(two_level_09_initial, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "JOYSCIE", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "COOPERATE", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "SC001Q01TA", hypothesis = NA) #not significant
waldTest(two_level_09_initial, type = "beta", coefs = "PROAT6", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "SCHSIZE", hypothesis = NA)
waldTest(two_level_09_initial, type = "beta", coefs = "SC048Q02NA", hypothesis = NA) #not significant, dropped
waldTest(two_level_09_initial, type = "beta", coefs = "SC048Q03NA", hypothesis = NA)



two_level_09_prefinal <- mix(PV1SCIE ~ ESCS+ ST062Q01TA+ ENVAWARE+ JOYSCIE + COOPERATE +
    USESCH +
                   + SC001Q01TA+PROAT6+ SCHSIZE+ SC048Q03NA+(1|CNTSCHID), data=df_9,
                      weights=c("fake_student_weights", "W_SCHGRNRABWT"))
summary(two_level_09_prefinal) #365.3 4666.9
(5714- 4666.9)/5714 #18,33%

waldTest(two_level_09_prefinal, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "JOYSCIE", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "COOPERATE", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "SC001Q01TA", hypothesis = NA) #not significant,
    dropped
waldTest(two_level_09_prefinal, type = "beta", coefs = "PROAT6", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "SCHSIZE", hypothesis = NA)
waldTest(two_level_09_prefinal, type = "beta", coefs = "SC048Q03NA", hypothesis = NA)



two_level_09_prefinal2 <- mix(PV1SCIE ~ ESCS+ ST062Q01TA+ ENVAWARE+ JOYSCIE + COOPERATE +
    USESCH +
                PROAT6+ SCHSIZE+SC048Q03NA+(1|CNTSCHID), data=df_9, weights=c("fake_
```

```
                           student_weights", "W_SCHGRNRABWT"))
summary(two_level_09_prefinal2) #361.9 #4677.1
(5714- 4677.1)/5714 #18.15%

waldTest(two_level_09_prefinal2, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "JOYSCIE", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "COOPERATE", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "PROAT6", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "SCHSIZE", hypothesis = NA)
waldTest(two_level_09_prefinal2, type = "beta", coefs = "SC048Q03NA", hypothesis = NA)




########
#10th graders
#########


#10 grade correlation matrix
df_10_cor<-data.frame(df_10$ESCS, df_10$ST062Q01TA, df_10$ENVAWARE, df_10$JOYSCIE, df_10$
    COOPERATE,df_10$USESCH,
                    df_10$SC001Q01TA,df_10$PROAT6,df_10$SCHSIZE,df_10$SC048Q02NA,df_10$
                        SC048Q03NA)
round(cor(df_10_cor),2)
corrplot(round(cor(df_10_cor, use="pairwise.complete.obs"),2))



#10 grade model
two_level_10_null <- mix(PV1SCIE ~ (1|CNTSCHID), data=df_10, weights=c("fake_student_weights", "W
    _SCHGRNRABWT"))
summary(two_level_10_null) #icc-18,06% 1373 6232

two_level_10_initial <- mix(PV1SCIE ~ ESCS+ ST062Q01TA+ ENVAWARE+ JOYSCIE + COOPERATE +
    USESCH +
                    SC001Q01TA +PROAT6+ SCHSIZE+SC048Q02NA+SC048Q03NA+(1|CNTSCHID), data=
                        df_10, weights=c("fake_student_weights", "W_SCHGRNRABWT"))
summary(two_level_10_initial) #174.5 4864.1
(6232-4864.1)/6232 # 21,95%
waldTest(two_level_10_initial, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_10_initial, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_10_initial, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_10_initial, type = "beta", coefs = "JOYSCIE", hypothesis = NA) #0.0553
waldTest(two_level_10_initial, type = "beta", coefs = "COOPERATE", hypothesis = NA)
waldTest(two_level_10_initial, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_10_initial, type = "beta", coefs = "SC001Q01TA", hypothesis = NA)
waldTest(two_level_10_initial, type = "beta", coefs = "PROAT6", hypothesis = NA) #0.0539
waldTest(two_level_10_initial, type = "beta", coefs = "SCHSIZE", hypothesis = NA)
waldTest(two_level_10_initial, type = "beta", coefs = "SC048Q02NA", hypothesis = NA) #0.1069
waldTest(two_level_10_initial, type = "beta", coefs = "SC048Q03NA", hypothesis = NA) #0.3818 dropped



two_level_10_prefinal1 <- mix(PV1SCIE ~ ESCS+ ST062Q01TA+ ENVAWARE+ JOYSCIE + COOPERATE +
```

```
   USESCH +
                           SC001Q01TA +PROAT6+ SCHSIZE+SC048Q02NA+(1|CNTSCHID), data=df_10,
                               weights=c("fake_student_weights", "W_SCHGRNRABWT"))
summary(two_level_10_prefinal1) #182.5 4863.4
(6232−4863.4)/6232 # 21,96%
waldTest(two_level_10_prefinal1, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "JOYSCIE", hypothesis = NA) #0.0638
waldTest(two_level_10_prefinal1, type = "beta", coefs = "COOPERATE", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "SC001Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "PROAT6", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "SCHSIZE", hypothesis = NA)
waldTest(two_level_10_prefinal1, type = "beta", coefs = "SC048Q02NA", hypothesis = NA) #0.1821 #dropped
   SC048Q02NA



two_level_10_prefinal2 <− mix(PV1SCIE ~ ESCS+ ST062Q01TA+ ENVAWARE+ JOYSCIE + COOPERATE +
   USESCH +
                           SC001Q01TA +PROAT6+ SCHSIZE+(1|CNTSCHID), data=df_10, weights=c("
                              fake_student_weights", "W_SCHGRNRABWT"))
summary(two_level_10_prefinal2) #187 4876
(6232−4876)/6232 # 21,76%
waldTest(two_level_10_prefinal2, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_10_prefinal2, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal2, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_10_prefinal2, type = "beta", coefs = "JOYSCIE", hypothesis = NA) #0.0691
waldTest(two_level_10_prefinal2, type = "beta", coefs = "COOPERATE", hypothesis = NA) #0.0541
waldTest(two_level_10_prefinal2, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_10_prefinal2, type = "beta", coefs = "SC001Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal2, type = "beta", coefs = "PROAT6", hypothesis = NA)
waldTest(two_level_10_prefinal2, type = "beta", coefs = "SCHSIZE", hypothesis = NA) #0.0812 dropped



two_level_10_prefinal3 <− mix(PV1SCIE ~ ESCS+ ST062Q01TA+ ENVAWARE+ JOYSCIE + COOPERATE +
   USESCH +
                           SC001Q01TA +PROAT6+(1|CNTSCHID), data=df_10, weights=c("fake_student
                               _weights", "W_SCHGRNRABWT"))
summary(two_level_10_prefinal3) #202.2 4887.3
(6232−4887.3)/6232 # 21,58%
waldTest(two_level_10_prefinal3, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_10_prefinal3, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal3, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_10_prefinal3, type = "beta", coefs = "JOYSCIE", hypothesis = NA) #0.0649 dropped
waldTest(two_level_10_prefinal3, type = "beta", coefs = "COOPERATE", hypothesis = NA) #0.0517
waldTest(two_level_10_prefinal3, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_10_prefinal3, type = "beta", coefs = "SC001Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal3, type = "beta", coefs = "PROAT6", hypothesis = NA)




two_level_10_prefinal4 <− mix(PV1SCIE ~ ESCS+ ST062Q01TA+ ENVAWARE + COOPERATE + USESCH +
                           SC001Q01TA +PROAT6+(1|CNTSCHID), data=df_10, weights=c("fake_student
```

```
_weights", "W_SCHGRNRABWT"))
summary(two_level_10_prefinal4) #183.6 4935.4
(6232−4935.4)/6232 # 20,81%
waldTest(two_level_10_prefinal4, type = "beta", coefs = "ESCS", hypothesis = NA)
waldTest(two_level_10_prefinal4, type = "beta", coefs = "ST062Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal4, type = "beta", coefs = "ENVAWARE", hypothesis = NA)
waldTest(two_level_10_prefinal4, type = "beta", coefs = "COOPERATE", hypothesis = NA)
waldTest(two_level_10_prefinal4, type = "beta", coefs = "USESCH", hypothesis = NA)
waldTest(two_level_10_prefinal4, type = "beta", coefs = "SC001Q01TA", hypothesis = NA)
waldTest(two_level_10_prefinal4, type = "beta", coefs = "PROAT6", hypothesis = NA)
```

References

[1] L. Araujo, A. Saltelli, and S. V. Schnepf. Do PISA data justify PISA-based education policy? *International Journal of Comparative Education and Development*, Vol. 19, No. 1, 2017.

[2] M. H. Bornstein and R. H. Bradley. *Socioeconomic Status, Parenting, and Child Development*. Routledge, 1 edition, 2002.

[3] S. J. Caldas and C. Bankston. Effect of School Population Socioeconomic Status on Individual Academic Achievement. *The Journal of Educational Research*, Vol. 90, No. 5, 1997.

[4] M. M. Chiu and L. Khoo. Effects of Resources, Inequality, and Privilege Bias on Achievement: Country, School, and Student Level Analyses. *American Educational Research Journal*, Vol. 42, No. 4, 2005.

[5] J. S. Coleman. Equality of Educational Opportunity. *Equity & Excellence in Education*, Vol. 6, No. 5, 1968.

[6] J. C. Hsu. Comparing the relationships between mathematics achievement and student characteristics in Canada and Hong Kong through HLM. 2007. Unpublished master's thesis, University of Victoria, Victoria, British Columbia, Canada.

[7] H. Huang. Can students themselves narrow the socioeconomic-status-based achievement gap through their own persistence and learning time? *Education Policy Analysis Archives*, Vol. 23, No. 108, 2015.

[8] M. Jackevičius. Putinaitė: kaimo mokyklose negali būti gerų rezultatų, kad ir kokie mokytojai ten bedirbtų. https://www.lrt.lt/naujienos/lietuvoje/2/1062665/putinaite-kaimo-mokyklose-negali-buti-geru-rezultatu-kad-ir-kokie-mokytojai-ten-bedirbtu, 2019. Accessed online 12-December-2022.

[9] A. Jaržemskis, L. Tamutienė, and S. Varnienė. *TIMSS tyrimo antrinė duomenų analizė: mokinių, jų tėvų ir mokytojų nuostatų ir lūkesčių kaita ir ryšys su mokinių pasiekimais*. Nacionalinė švietimo agentūra, 2022.

[10] J. Mang, H. Küchenhoff, S. Meinck, and M. Prenzel. Sampling weights in multilevel modelling: an investigation using PISA sampling structures. *Large-scale Assessments in Education*, Vol. 9, No. 6, 2021.

[11] K. Marjoribanks. *Families and their Learning Environments: An Empirical Analysis*. Routledge, 1 edition, 1979.

[12] Y. Mo, K. Singh, and M. Chang. Opportunity to learn and student engagement: a HLM study on eighth grade science achievement. *Educational Research for Policy and Practice*, Vol. 12, 2013.

[13] A. Murauskaitė. Išreitingavo švietimo bėdas, skaudžiausios – kokybės skirtumai ir gerų mokytojų stygius. https://www.lrt.lt/naujienos/lietuvoje/2/1244605/isreitingavo-svietimo-bedas-skaudziausios-kokybes-skirtumai-ir-geru-mokytoju-stygius, 2020. Accessed online 12-December-2022.

[14] S. Noël and P. de Broucker. Intergenerational Inequities: A comparative analysis of the influence of parents' educational background on length of schooling and literacy skills. *In Pursuit of Equity in Education Using International Indicators to Compare Equity Policies*, pages 277–298, 2001.

[15] OECD. *Learning for Tomorrow's World: First Results from PISA 2003*. OECD Publishing, Paris, 2004.

[16] OECD. In *PISA Data Analysis Manual SPSS, SECOND EDITION*, pages 93–101. OECD Publishing, Paris, 2009.

[17] OECD. PISA 2009 Results: Executive Summary. 2010.

[18] OECD. Context Questionnaire Development. https://www.oecd.org/pisa/data/PISA-2015-Technical-Report-Chapter-3-Context-Questionnaire-Development.pdf, 2017. Accessed online 12-December-2022.

[19] OECD. Procedures and Construct Validation of Context Questionnaire Data. https://www.oecd.org/pisa/sitedocument/PISA-2015-Technical-Report-Chapter-16-Procedures-and-Construct-Validation-of-Context-Questionnaire-Data.pdf, 2017. Accessed online 12-December-2022.

[20] OECD. Programme for International Student Assessment: an overview. https://www.oecd.org/pisa/data/2015-technical-report/01_Chapter_01_PISA2015.pdf, 2017. Accessed online 12-December-2022.

[21] OECD. Scaling PISA Data. https://www.oecd.org/pisa/sitedocument/PISA-2015-Technical-Report-Chapter-9-Scaling-PISA-Data.pdf, 2017. Accessed online 12-December-2022.

[22] OECD. Survey Weighting. https://www.oecd.org/pisa/data/2015-technical-report/PISA-2015-Technical-Report-Chapter-8-Survey-Weighting.pdf, 2017. Accessed online 12-December-2022.

[23] OECD. Test Design and Development. https://www.oecd.org/pisa/data/2015-technical-report/PISA-2015-Technical-Report-Chapter-2-Test-Design-and-Development.pdf, 2017. Accessed online 12-December-2022.

[24] OECD. In *PISA 2018 Results (Volume II): Where All Students Can Succeed*, pages 15–28. OECD Publishing, Paris, 2019.

[25] OECD. In *PISA 2018 Results (Volume II): Where All Students Can Succeed*, pages 49–62. OECD Publishing, Paris, 2019.

[26] J. W. Osborne. Advantages of Hierarchical Linear Modeling. *Practical Assessment, Research, and Evaluation*, Vol. 7, No. 1, 2000.

[27] L. Perry and A. McConney. School Socio-Economic Composition and Student Outcomes in Australia: Implications for Educational Policy. *Australian Journal of Education*, Vol. 54, No. 1, 2010.

[28] H. Rindermann and A. E. Baumeister. Parents' SES vs. parental educational behavior and children's development: A reanalysis of the Hart and Risley study. *Learning and Individual Differences*, Vol. 37, 2015.

[29] J. B. Schreiber and B. W. Griffin. Review of Multilevel Modeling and Multilevel Studies in The Journal of Educational Research (1992-2002). *The Journal of Educational Research*, Vol. 98, No. 1, Pages 24-34, 2004.

[30] S. R. Sirin. Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, Vol. 75, No. 3, 2005.

[31] W. F. Tate. Race-Ethnicity, SES, Gender, and Language Proficiency Trends in Mathematics Achievement: An Update. *Journal for Research in Mathematics Education*, Vol. 28, No. 6, 1997.

[32] A. Zabulionis. *Tarptautinio švietimo tyrimo OECD PISA Lietuvos ir kaimyninių šalių duomenų tikslinė antrinė analizė*. Nacionalinė švietimo agentūra, 2020.