

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Abrigo, Michael Ralph M.; Lingatong, Edmar E.; Relos, Charlotte Marjorie L.

Working Paper School bullying contributes to lower pisa achievement among Filipino students: Who gets bullied? Why does it matter?

PIDS Discussion Paper Series, No. 2024-45

Provided in Cooperation with: Philippine Institute for Development Studies (PIDS), Philippines

Suggested Citation: Abrigo, Michael Ralph M.; Lingatong, Edmar E.; Relos, Charlotte Marjorie L. (2024) : School bullying contributes to lower pisa achievement among Filipino students: Who gets bullied? Why does it matter?, PIDS Discussion Paper Series, No. 2024-45, Philippine Institute for Development Studies (PIDS), Quezon City, https://doi.org/10.62986/dp2024.45

This Version is available at: https://hdl.handle.net/10419/316143

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

DISCUSSION PAPER SERIES NO. 2024-45

School Bullying Contributes to Lower PISA Achievement among Filipino Students: Who Gets Bullied? Why Does It Matter?

Michael R.M. Abrigo, Edmar E. Lingatong, and Charlotte Marjorie L. Relos



Philippine Institute for Development Studies

The PIDS Discussion Paper Series constitutes studies that are preliminary and subject to further revisions. They are being circulated in a limited number of copies only for purposes of soliciting comments and suggestions for further refinements. The studies under the Series are unedited and unreviewed. The views and opinions expressed are those of the author(s) and do not necessarily reflect those of the Institute. Please seek approval from the authors before citing this publication.

This study is carried out with support from the Second Congressional Commission on Education (EDCOM II).



CONTACT US: RESEARCH INFORMATION DEPARTMENT Philippine Institute for Development Studies School Bullying Contributes to Lower PISA Achievement among Filipino Students: Who Gets Bullied? Why Does It Matter?

> Michael R.M. Abrigo Edmar E. Lingatong Charlotte Marjorie L. Relos

PHILIPPINE INSTITUTE FOR DEVELOPMENT STUDIES

December 2024

Abstract

The Philippines has one of the highest school bullying rates around the world. While its nature, causes and impacts are well-documented in the international literature, the local evidence remains limited and at times conflicting. In this study, we assess the contribution of bullying exposure on student achievement in an international large-scale student assessment, and infer its potential long-term implications. We find that differences in bullying exposure explains around 0.05 standard deviation of the gap in average student achievement between proficient and non-proficient students, which could potentially cost the Philippines around PhP10-20 billion annually in foregone economic activity. We also document important risk factors of being the "most bullied" around the world, and highlight spatial disparities in bullying risks.

Keywords: School bullying, Basic education, Learning loss, Learning achievement, Small area estimate, Machine learning

Table of Contents

1.	Introduction	1
2.	What does bullying look like?	3
3.	Does bullying matter?	5
4.	Who gets bullied?	10
5.	Where is bullying risk highest?	14
6.	Policy implications	17
Refe	erences	18
App	endix A. Blinder-Oaxaca decomposition	22
App	endix B. Small area estimation of bullying risk	36
App	endix C. Robustness analysis	39

List of Tables

Table 1. Share (%) of students who experienced bullying	4
Table 2. Log-odds of being among world's "most bullied" students	13

List of Figures

Figure 1. Bullying exposure score average and dispersion by country	3
Figure 2. Blinder-Oaxaca decomposition: Endowment effects	8
Figure 3. Blinder-Oaxaca decomposition: Endowment and Unexplained effects	9
Figure 4. Predicted v. actual "most bullied" bullying rates by school	11
Figure 5. Covariate influence	12
Figure 6. Risk of being among "most bullied" students by province	15
Figure 7. Trimmed range by province	16

School Bullying Contributes to Lower PISA Achievement among Filipino Students: Who Gets Bullied? Why Does It Matter?

Michael R.M. Abrigo⁽²⁾, Edmar E. Lingatong, and Charlotte Marjorie L. Relos¹

1. Introduction

The Philippines has one of the highest school bullying rates around the world (OECD, 2023a; Mullis, et al., 2020). While its nature, causes and impacts have been well documented in the international literature, particularly in developed countries, the nascent local empirical literature remains limited.

School bullying has far-reaching implications. Victims suffer from poorer mental health (Moore, et al., 2017; Schoeler, 2018), are more likely to experience psychosomatic conditions (Gini, et al, 2014; Gini and Pozzoli, 2009), and have worse academic outcomes (Halliday, et al., 2021). It haunts its victims into adulthood with its long-term negative impacts on labor force participation, employment, income and wealth accumulation (Brimblecombe, et al., 2018; Drydakis, 2014; Mukerjee, 2018; Sansone, Leung and Wiederman, 2012).

Evidence from observational studies in the Philippines are mixed. Orbeta, et al. (2020), for example, found strong consistent negative association between bullying exposure and test scores in an international large-scale student assessment (ILSA). However, results in Bernardo, et al. (2023) using the same ILSA show some protective effects of bullying exposure against low science achievement. Herrin, et al. (2024), using a nationally representative cohort of children, also document promotive effects of bullying on on-time school attendance. This mixed bag warrants further inquiry to guide education policy in the country.

In this study, we re-examined the association between students' bullying exposure and their achievement in ILSA using the 2022 Programme for International Student Assessment (PISA) by the Organization for Economic Co-operation and Development (OECD, 2023a). Unlike previous studies that aimed to identify correlates of student achievement, our goal is to estimate the magnitude of contributions of these factors. This allows us to provide some indicative measure of learning losses from bullying exposure, rank it against other potential sources of learning disparities, and ultimately put a price tag on foregone opportunities because of students' bullying exposure in schools.

We then take a step back to describe what bullying looks like in the Philippines, and to identify factors that correlates with bullying victimization. We focus on the 2022 PISA's "most bullied" students, representing the top ten percentile of students in 80 countries with the highest bullying exposure score. We used classical and state-of-the-art machine learning techniques to predict a Filipino student's risk of being tagged as part of the "most bullied" students globally. We select the best model to predict school-level bullying risk, which we aggregate into provincial small area estimates of "most bullied" bullying risk.

The results of our analysis using Blinder-Oaxaca decomposition suggest that school bullying have significant and measurable negative association with learning achievements across

¹ Fellow II, Project Technical Specialist, and Project Technical Specialist, respectively, at the Philippine Institute for Development Studies.

^A Corresponding author

subjects, capturing about five PISA points of the gap in average scores between proficient and non-proficient students. While the magnitude of potential learning losses from bullying appears small, this translates to about 0.05- to 0.08-percentage point reduction in long-term economic activity, or roughly a foregone PhP10-20 billion per year. This highlights the non-trivial cost of school bullying in the economy, and the necessity to develop focused large-scale programs addressing its causes and dampening its negative impacts.

The Philippine government has made important strides in enshrining policies to protect children from various forms of abuses. In 1992, the Philippines enacted Republic Act (RA) 7610, which provides special protection to children from acts or conditions that may be detrimental to their development, including psychological and physical abuse and emotional maltreatment. Department of Education Order No. 40, s. 2012, provides for a national child protection policy in schools, outlining the role of different stakeholders in preventing, addressing and handling child abuse, exploitation, violence and discrimination cases, including bullying or peer abuse, in schools. The adoption of RA 10627 or the Anti-Bullying Act of 2013 required basic education institutions to adopt policies to prevent and address school bullying.

Despite these measures, the 2022 PISA show that about three in every four students have had experienced bullying acts at least once in the past 12 months. More than a third of students have had experienced at least one form of bullying every week. Among the "most bullied" Filipino students, about two percent experience at least nine acts of bullying regularly.

The most important correlates of bullying we identified in this study aligns with global experiences. Disadvantaged children, particularly those who have had experienced hunger or are grade repeaters, have higher risks of being a "most bullied" student. Attending a private school or being female offers some protective effects against bullying. These offer policymakers and administrators some guidance in designing both preventive and protective measures against school bullying. These also hints on the complex nature of school bullying, touching on individual, household and community factors, that may be beyond the purview of the education system.

Provincial small area estimates of bullying risk highlight several potential areas of concern across the country. Areas with high average but low dispersion of bullying risks across schools within provinces may need immediate attention from administrators as students are similarly exposed to high bullying risks.

The rest of the study is organized as follows. In the next section, we describe what school bullying looks like in the Philippines, focusing on the 2022 PISA's "most bullied" students around the world. This is followed in Section 3 with our Blinder-Oaxaca decomposition of the gap in average PISA scores of proficient and non-proficient students into its contributing factors. In Section 4, we take a step back to identify correlates of bullying exposure, which we then use to generate provincial small area estimates of bullying risks that we present in Section 5. Finally, in Section 6, we conclude with a summary and some implications for policy.

2. What does bullying look like?

Comparative cross-country estimates from the 2022 PISA (OECD, 2023a) show that the Philippines has the highest average bullying exposure score among 15-year-old students in 80 participating countries. This score is based on nine questions that asked students about how frequently they experienced bullying acts in the past 12 months prior to the PISA test, including (1) being left out on purpose, (2) being made fun of, (3) being threatened, (4) having other students take away or destroy personal belongings, (5) being hit or pushed around, (6) being the subject of nasty rumors, (7) being on a physical fight on school property, (8) having stayed home because of feeling unsafe, and (9) being extorted for money at school. Individual responses in these items were transformed into a single bullying experience score using a generalized partial credit model, with the resulting scores standardized relative to the distribution of bullying experience scores in OECD countries in the 2018 PISA.

Bullying experiences among 15-year-old students in the Philippines are relatively more varied compared with same-aged students in other countries as shown in Figure 1. The plot shows the average bullying exposure score against its standard deviation across PISA 2022 participating countries, where the Philippines is markedly distinguishable as it bounds both measures from above. Further analysis of the 2022 PISA student-level data shows that while there were 30% of Filipino students with bullying exposures scores below the 2018 OECD average – i.e., they experienced less bullying relative to an average student from OECD countries in 2018 – there were, however, 36% of Filipino students in the same international student assessment who were in the top ten percent of most bullied students across all participating countries.



Figure 1. Bullying exposure score average and dispersion by country

Source: Authors' estimates based on PISA 2022 (OECD, 2023b). Note: Each point on the plot refers to one of 80 countries in the 2022 PISA. IDN – Indonesia, JPN – Japan, KHM – Cambodia, KOR – South Korea, MYS - Malaysia, SGP – Singapore, THA – Thailand, PHL – Philippines.

Bullying experience is rather pervasive among Filipino students. But those who were in the 2022 PISA's "most bullied" students globally were more likely to have had experienced any and all of the nine bullying acts documented by the international assessment as shown in Table 1. All "most bullied" Filipino students have had experienced at least one bullying act in the past 12 months prior to the PISA 2022, compared with 76.6% of all Filipino students in the assessment. More than a tenth of the "most bullied" Filipino students have had experienced all nine bullying acts in the past year prior to the assessment, while this only happened to less than four percent of all Filipino students.

The "most bullied" Filipino students were also exposed to bullying acts more frequently. About three in every four "most bullied" Filipino students have had experienced at least one bullying act at least once a week in the past 12 months, compared with only a third of all Filipino students. The most common and frequent form of bullying experienced by the "most bullied" Filipino students were verbal ("Other students made fun of me", "I was threatened by other students"), although physical ("I got hit or pushed around by other students", "Other students took away or destroyed things that belonged to me") and relational ("Other students left me out of things on purpose", "Other students spread nasty rumors about me") forms of bullying were also quite common.

While only 0.6% of all Filipino students in PISA 2022 experienced all nine bullying acts at least once a week in the past 12 months, extrapolating this to all 15-year-old students in school year 2021-2022 translates to at least 11,000 students who were regularly bullied in many and varied forms across the country.

	At leas in past 12	t once months	Once a week or more in past 12 months		
	All	Most	All	Most	
	students	bullied	students	bullied	
Other students left me out of things on purpose.	34.8	62.1	7	17.1	
Other students made fun of me.	62	90.9	16.3	35.7	
I was threatened by other students.	41.2	79.1	8.9	23.1	
Other students took away or destroyed things that belonged to me.	32.8	67.8	6.7	18.1	
l got hit or pushed around by other students.	38.2	71.7	7.5	19.7	
Other students spread nasty rumors about me.	33.7	66.7	7	18.2	
I was on a physical fight on school property.	27	55.1	5.3	13.6	
l stayed home because l felt unsafe.	26.9	56.2	6.5	15.7	
I gave money to someone at school because they threatened me.	25.2	56.8	7	17.7	
Experienced at least one of listed activities	76.6	100	34.5	75.5	
Experienced all listed activities	3.9	10.8	0.6	1.6	

Table 1. Share (%) of students who experienced bullying

Source: Authors' estimates based on PISA 2022 (OECD, 2023b). Note: "Most bullied" students refer to Filipino students who were in the top ten percentile of students globally in the 2022 PISA with the highest bullying exposure score.

3. Does bullying matter?

The impact of school bullying on various academic, psychological, health and social outcomes have been well-documented in the international literature. Several meta-analyses, for example, have found that bullying has strong causal association with mental health problems (Moore, et al., 2017; Schoeler, 2018), including depression, anxiety, poor general health, and suicidal ideation and behaviors, with females suffering from worse outcomes compared with males (Halliday, et al., 2021). Bullied children and adolescents have higher risk for headaches (Gini, et al., 2014) and other psychosomatic problems (Gini and Pozzoli, 2009). They have poorer immediate academic outcomes and school connectedness, which may last for extended periods (Halliday, et al., 2021). They also tend to exhibit anti-school behaviors, such as substance use, isolation, violence, and truancy (Pengpid and Peltzer, 2019). Victims, bullies and bully-victims were also documented to be more likely to carry weapons compared with uninvolved peers (van Geel, Vedder and Tanilon, 2014). Bullying victimization has been shown to have long-term negative impacts on labor force participation, employment, income and wealth accumulation (Brimblecombe, et al., 2018; Drydakis, 2014; Mukerjee, 2018; Sansone, Leung and Wiederman, 2012).

The limited available evidence in the Philippines appear to be mixed. Chiu and Vargo (2022), using the Philippine round of the 2011 Global School-based Health Survey, found that bullying victimization is associated with higher risk of having suicidal ideation and ever attempting suicide, although it appears to play only a secondary role compared to other variables explored in their analysis. Orbeta, et al., (2020), using PISA 2018 data, documented consistent negative association of bullying and test scores in mathematics, science and reading. However, Bernardo, et al. (2023), analyzing the same PISA 2018 data, showed that bullying is negatively associated with a child having low proficiency in science. Furthermore, Herrin, et al. (2024), using survey responses from the nationally representative panel of children in the Longitudinal Cohort Study on the Filipino Children, found positive association with some forms of bullying and on-time school attendance, pointing to protective effects of some human capital factors, such as child's IQ and parents' education.

We re-examined the contribution of bullying exposure on academic performance in mathematics, science and reading explored earlier by Orbeta, et al. (2020) and Bernardo, et al. (2023), but using student-level assessment results in the 2022 PISA in the Philippines. We decomposed the difference in average assessment scores of proficient and non-proficient students into contributions of different academic performance predictors using Blinder-Oaxaca decomposition (Blinder, 1974; Oaxaca, 1974). In the analysis, the difference in average scores of the two groups are decomposed as due to differences in average characteristics, called explained or endowment effects, and as due to differences in how the groups respond to changes in these characteristics, called the unexplained or coefficient effects. Further details of the statistical method employed are provided in Appendix A.

Similar to Orbeta, et al. (2020) and Bernardo, et al. (2023), we also used student responses in the PISA student background questionnaire. This allowed us to assess the contribution of student-level variables, such as student's motivation and beliefs, school experiences and home background. We supplement these variables with school-level indicators based on school administrative data collected by the Department of Education (DepEd) and by community-level characteristics derived from the 2020 Census of Population and Housing (CPH) by the Philippine Statistics Authority (PSA). Community-level data on conflict and nighttime lights taken from the Uppsala Conflict Data Program (UCDP) (Davies, et al., 2024) and the Earth

Observation Group (EOG) (Elvidge, et al., 2017), respectively, were also included to proxy for social factors that may affect student achievement. Further information on variables included in the model is provided in Appendix A.

However, unlike Bernardo, et al. (2023) and Orbeta, et al. (2020) that both aimed to identify background factors associated with student proficiency, our goal is to measure how large the contribution of such individual, school and community factors are on differences in average test scores of proficient and non-proficient students. As part of the analysis, we also document in this section the disparities in these endowments among proficient and non-proficient students, which could point to potential avenues for future policy intervention.

The Philippines has one of the lowest average scores in PISA 2022. This is despite some modest but measurable gains from the PISA 2018 (OECD, 2023c) and the documented learning losses from school closures due to the 2020-2023 global pandemic elsewhere (e.g., Jacubowski, Gajderowicz and Patrinos, 2023). Assessment scores in PISA are normed to the baseline OECD distribution in 2000 with scaled mean of 500 points and standard deviation of 100 points. An average Filipino student in the PISA 2022 scored 355 points in mathematics, 347 points in reading, and 356 points in science. That is about 1.5 standard deviations lower relative to an average student in OECD countries. Majority of Filipino students in PISA 2022 were ranked as having only level 1 proficiency, i.e., the lowest in the PISA proficiency scale of 1 to 6, across tested subjects: mathematics (84.0%), reading (76.3%), and science (93.2%).

The difference in average scores between proficient students, defined in our analysis as those having proficiency level 2 or higher, and non-proficient students, defined as those in level 1, are quite substantial. The gap is widest in reading (154.1 points), followed by science (140.8 points), and finally mathematics (121.9 points). It is important to point out, however, that students with low PISA 2022 achievement were also likely to be disadvantaged in other dimensions. They are more likely to be bullied, have had missed school, or repeat grade levels. They are also more likely to attend schools with greater shortage of educational materials and have lower capacity to provide remote instruction. They are more likely to come from households with lower socioeconomic status.

Figure 2 summarizes these differences in endowments between proficient and non-proficient students and its contribution to disparities in average scores in PISA 2022. In each panel, the horizontal axis provides the direction of the disparity in average endowments, with values to the right indicate higher average endowments for proficient students while those to the left favor non-proficient students. The vertical axis, on the other hand, shows the contribution of the differences in endowments on the disparity in PISA scores between proficient and non-proficient students. Values above the horizontal line indicate positive contributions to average achievement scores of proficient students, while those below indicate positive contributions to non-proficient students.

Differences in socio-economic status (SES) contribute the largest gap in achievement scores between proficient and non-proficient students across tested subjects in PISA 2022, explaining as much as 9 points of the observed difference in scores. Missing school for more than 3 months (MISSSCH), grade repetition (REPEAT) and bullying exposure (BULLIED) — all higher on average among non-proficient students — each explain between 3 and 7 points of the observed PISA disadvantage by non-proficient students. Other student factors, such as curiosity (CURIOUS), expected level of education (EXPECEDU) and working for pay (WORKPAY), as well as school environment factors like the school's safety risks (SCHRISK) and students'

age dispersion (AGEVAR) also contribute significantly to the observed disparity in PISA outcomes between proficient and non-proficient students. A full list of variables and endowment effects estimates are provided in Appendix A.

While socioeconomic status (SES) remains the largest contributing factor in disparity in achievement scores between proficient and non-proficient students, many of the other prominent variables are either school environment factors (BULLIED, SCHRISK, AGEVAR) or student-specific factors (MISSSCH, REPEAT, CURIOUS) that may be more readily modifiable in the near-term than socioeconomic status. These specific school environment and student-level variables together accounts for about 19- to 30-PISA points of the gap between the average achievement scores of proficient and non-proficient students across tested subjects.

The endowment effects shown in Figure 2 capture only about a third of the observed difference in average scores of proficient and non-proficient students, while the rest remain unexplained. This suggests that differences in how students in the two groups respond to endowments, i.e., coefficient effects, also matter greatly. But for bullying exposure, this appears to be of limited importance as estimated unexplained effects across tested subjects are all close to zero, indicating that proficient and non-proficient students respond quite similarly on average to bullying exposure in school.

Figure 3 displays the top ten variables with highest endowment effects and their corresponding unexplained effects in each subject area. Most of the factors affecting achievement scores show negligible unexplained effects, suggesting that the differences in average levels of these factors drive the disparity in average achievement scores between the proficient and non-proficient groups. There are variables, however, such as internet connectivity (INTERNET) and old-age dependency ratio (DEPOLD) that show a considerable amount of unexplained effect, hinting that proficient and non-proficient students respond to these endowments differently.

In science, for example, internet connectivity (INTERNET) and socioeconomic status (SES) similarly captures about nine PISA points each of the difference in average scores between proficient and non-proficient students. However, the endowment effect for internet connectivity is relatively much smaller compared with that for socioeconomic status. How students take advantage of their school's internet connectivity matters more than connectivity itself. As for bullying, the bullying experience itself (i.e. regardless of how the students respond to having been bullied) appears to explain most of its contribution towards the gap in achievement scores across all subject areas.

To put the above estimates in perspective, we follow Jacubowski, Gajderowicz and Patrinos (2023) by estimating the potential foregone economic activity from learning losses but in our case due to bullying instead of school closures. Using a factor of 1.74 estimated by Hanushek and Woessmann (2010) to quantify the response of a country's gross domestic product (GDP) growth to a standard deviation increase in average cognitive skills, and accounting for the share of non-proficient Filipino students in the 2022 PISA and the estimated endowment effects of bullying, our back of the envelope calculation points to school bullying costing the Philippines between 0.05- to 0.08-percentage points in GDP growth because of lower school achievement. When applied to the Philippine GDP of PhP24.3 trillion in 2023, our estimates suggest a foregone PhP10-20 billion in annual GDP due to school bullying.



Figure 2. Blinder-Oaxaca decomposition: Endowment effects

Source: Authors' calculations based on 2022 PISA data (OECD, 2023b). Note: The figure plots the standardized mean difference and the endowment effect of each of the independent variables of the Oaxaca-Blinder Decomposition. Variables with endowment effects that are greater than 1 or less than -1 are highlighted and labeled.



Figure 3. Blinder-Oaxaca decomposition: Endowment and Unexplained effects

Source: Authors' calculations based on 2022 PISA data (OECD, 2023b). Note: The figure shows the top ten variables with highest endowment effects and their corresponding unexplained effects based on Oaxaca-Blinder Decomposition. See Table A1.1 for a full description of the variables. See also Table A1.3 for the full list of endowment and unexplained effects of each variable.

4. Who gets bullied?

Given the substantial contribution of school bullying to learning loss, as seen from the Oaxaca-Blinder decomposition in the previous section, it is instructive to examine important risk factors of bullying victimization in schools. In order to identify these, we specify several binary classification models that aim to predict a Filipino student's probability of being part of the 2022 PISA's "most bullied" students. More specifically, we used simple logistic regression and two common machine learning techniques, i.e., lasso (least absolute shrinkage and selection operator) and elastic net regressions, that build on the simple logistic regression model by performing variable selection and regularization to improve prediction accuracy. We also tuned a gradient-boosted classifier (GBC) that combines multiple weak logistic regression trees to create a more accurate classifier. Unlike in the first three models wherein functional forms of explanatory variables are specified a priori, the GBC learns such relationships from the data. Further details are provided in Appendix B.

We include student, school, and community characteristics as explanatory variables. These variables were selected based on several considerations, including their potential predictive ability as documented in the literature, their coverage to allow the greatest number of observations to be included in the models, and the availability of school-level analogues in the case of individual-level variables to allow out-of-sample predictions to non-PISA schools. Similar to the approach in the Oaxaca-Blinder decomposition in the previous section, student-level characteristics from the 2022 PISA were supplemented by school-level variables from DepEd administrative data, and by community-level characteristics from the UCDP and EOG, and those derived from the 2020 CPH by the PSA.

All the classifiers show acceptable predictive accuracy with their areas under the receiver operating characteristic curve (AUC) being greater than 0.7 (c.f. Hosmer and Lemshow, 2000). The AUC, which ranges between 0 and 1, provides a summary measure of how well a binary classifier discriminates between class types, in this case, "most bullied" or otherwise. An AUC of 0.5 indicates that a classifier performs as well as random guesses, whereas higher values indicate better discrimination. While the calculated AUC of the GBC (0.749) may be qualitatively superior to those of the simple (0.730), lasso (0.727) and elastic net (0.729) logistic regressions, their 95% confidence intervals are overlapping.

As an additional goodness-of-fit check, we predict the share of "most bullied" students by school based on the above model using school-level averages of the explanatory variables, which we compared with actual school-level shares calculated directly from the 2022 PISA. The results presented in Figure 4 shows high positive correlation between predicted and actual "most bullied" rates by school, although the GBC exhibit quite substantial drift relative to actual rates. This is confirmed by a linear regression of actual rates on the predicted share of "most bullied" students by school that show statistically significant departure from the expectation of unit-slope and zero-intercept for the GBC.

The estimated coefficients of determination of regression lines in Figure 4 range between 0.699 for the lasso model and 0.847 for the GBC (see Appendix B for details), suggesting that our classifiers capture substantial portion of between-school variation in the proportion of students belonging in the 2022 PISA's "most bullied" students globally. However, it also points to the complex nature of bullying victimization as our classifiers fail to capture up to 30 percent of the observed variation in "most bullied" rates across schools.

Using GBC influence statistics, we then characterize the importance of explanatory variables in predicting a student being part of the world's "most bullied" students. The influence statistic captures the share of the log-likelihood of the regression tree model explained by a particular variable. A split on a variable in the regression tree increases the model's log-likelihood. In a linear regression model with Gaussian errors, this is proportional to an increase in the sum of squares explained by the model. The influence statistics are presented in Figure 5.



Figure 4. Predicted v. actual "most bullied" bullying rates by schoolA. Simple logistic regressionB. LASSO logistic regression

Source: Authors' calculations based on 2022 PISA data (OECD, 2023b). Note: The figure plots the estimated share of students that belong to 2022 PISA's "most bullied" students globally against predicted rates based on the specified model. The predicted rates are estimated using school-level analogues of student-level characteristics. See Appendix B for details.

Based on the estimated influence statistics, the most important factors in explaining the probability of belonging among the "most bullied" students include (i) experiencing hunger in the past month (23.4%), (ii) being a grade repeater (19.8%), (iii) studying in a public school (10.5%), (iv) municipality-level dependency ratio (8.9%), and (v) being female (4.1%). These five factors together capture as much as two-thirds of the log-likelihood of the GBC.

We also explored the potential influence of having a school principal as school head (instead of a head teacher) and of employing a guidance counselor in the school on the risk of a student being part of the world's "most bullied". However, we excluded private schools in this analysis as such information in the DepEd database is only available for public schools. The results suggest that having a school principal as school head or employing a guidance counselor have limited influence on bullying victimization risk. See Appendix C for details.



Figure 5. Covariate influence

Source: Authors' calculations based on 2022 PISA data (OECD, 2023b). Note: The influence statistic captures an explanatory variable's share of the log-likelihood of a regression tree model in a GBC. A higher share signifies greater importance in explaining the model.

	(1)	(2)	(3)	(4)
Grade repeater (=1)	0.960***	0.848***	0.812***	0.812***
	(0.073)	(0.072)	(0.073)	(0.073)
Female (=1)	-0.376***	-0.380***	-0.387***	-0.387***
	(0.055)	(0.053)	(0.053)	(0.053)
Transferee student (=1)	0.299***	0.201***	0.189***	0.186***
	(0.067)	(0.066)	(0.067)	(0.067)
Experienced hunger in past month (=1)	0.887***	0.835***	0.820***	0.820***
	(0.058)	(0.058)	(0.058)	(0.059)
Public school (=1)		0.730***	0.744***	0.746***
		(0.150)	(0.136)	(0.132)
Total enrollment, ln		-0.165***	-0.065	-0.053
		(0.042)	(0.045)	(0.047)
Share of females in enrolled, G10		-1.133**	-1.682***	-1.722***
		(0.555)	(0.579)	(0.565)
Share of repeaters in enrolled, G10		-4.460	0.430	-1.569
		(4.060)	(5.309)	(5.138)
Age, school standard deviation		0.020	0.006	-0.023
		(0.094)	(0.087)	(0.092)
Age, G10 standard deviation		0.008	-0.139	-0.136
		(0.172)	(0.166)	(0.166)
Share of transferees in enrolled, G10		-1.489	2.501	2.863
		(2.654)	(2.646)	(2.581)
Share of 4Ps beneficiaries in enrolled, G10		0.390*	0.185	0.193
		(0.210)	(0.218)	(0.217)
Share of IPEd students in enrolled, G10		0.366**	0.244	0.154
		(0.159)	(0.204)	(0.206)
School offers elementary level (=1)		-0.431	-0.373	-0.263
		(0.285)	(0.323)	(0.322)
School offers senior high school level (=1)		0.243**	0.106	0.082
		(0.108)	(0.097)	(0.101)
Share of aged 20+ years in enrolled		-0.039	-0.003	0.003
		(0.029)	(0.035)	(0.035)
Constant	-1.080***	-0.075	-7.673**	-6.690*
	(0.061)	(0.421)	(3.492)	(3.803)
With community variables			Yes	Yes
With location fixed effects				Yes
Pseudo R-sq.	0.097	0.115	0.122	0.122
BIC	8,058	7,998	8,030	8,061
Ν	6,782	6,782	6,782	6,782

Table 2. Log-odds of being among world's "most bullied" students

Source: Authors' calculations based on 2022 PISA data (OECD, 2023b). Note: The model predicts the probability of a Filipino student being part of the world's "most bullied" students using simple (binary) logistic regression. Values in parentheses are heteroskedasticity-robust standard errors clustered at the school level. *, **, *** denote statistical significance at the 10-, 5- and 1-percent alpha-levels, respectively. BIC – Bayesian information criterion; N – number of observations.

While the influence statistic captures the variation explained by a particular variable, it cannot distinguish the direction of the relationship between an explanatory variable and the outcome. The functional form of the relationship between outcome and an explanatory variable is typically explored through visualization, wherein outcomes are predicted for specific values of explanatory variables. We depart from this tradition by instead relying on parameter estimates from our simple logistic regression model.

The simple logistic regression estimates presented in Table 2 show the (conditional) log-odds of a Filipino student being part of the world's "most bullied" students. Each column represents different sets of explanatory variables. We start with the most basic model, where we include only student-level characteristics, and then sequentially add school- and community-level characteristics, as well as location fixed effects, to assess the robustness of the estimates to the inclusion of previously excluded controls. Among the specified models in Table 2, the second model has the lowest Bayesian information criterion and is thus selected as the preferred model.

Focusing on the most influential student- and school-level variables identified by the GBC, results from the simple logistic regression suggest that students who experienced hunger in the past week, are grade repeaters, or are enrolled in a public school have higher risks of being among the "most bullied" students. Having any of these characteristics at least doubles the odds of being among the "most bullied". Being female, on the other hand, provides some protective effects against being among the "most bullied" students, with females being 31.6% less likely to be in the "most bullied" group compared with males of comparable characteristics. These results are generally in line with associations found in the literature (e.g. Carbone-Lopez, Esbensen, and Brick, 2010; Chen, et al., 2024; Jeong, et al., 2013; Lian, et al., 2021; Hosozawa, et al., 2021).

5. Where is bullying risk highest?

Using the classifiers in the previous section, we then predict the risk of a student to be included in the world's "most bullied" group in each junior high school in the DepEd database. We focus on the predicted bullying risks from the simple logistic regression model, given its performance discussed in the previous section. In this exercise, we replaced individual-level variables with school-level averages in order to provide a first-order approximation of school-level risks. In particular, we replaced the individual-level indicator variables listed in Table 2 with school- or community-level shares, e.g., female student indicator with share of females in enrolled, repeater indicator with repeater ratio, etc. Further details are discussed in Appendix B.

These estimates are then aggregated at the province level to match the highest level of aggregation of explanatory variables employed in the classifiers. Separate estimates are provided for schools in the public and private sectors. These small-area "most bullied" bullying risk estimates are summarized as heat maps in Figure 6. We also mapped the trimmed range, covering the 5th and 95th percentiles of school-level estimates, to provide some indication of the dispersion of bullying risks across schools within provinces. These are shown in Figure 7.

Figure 6. Risk of being among "most bullied" students by province



Source: Authors' estimates. Note: The plot maps the provincial average risk of a student to be part of PISA 2022's "most bullied" students around the world. Expected risks are estimated for each junior high school and aggregated at the province level. See Appendix B for estimation details.



Source: Authors' estimates. Note: The plot maps the provincial 5th v. 95th percentile trimmed range of expected risk of a student to be part of PISA 2022's "most bullied" students around the world. See Appendix B for estimation details.

Several observations from the small-area estimates are worth noting. First, the risk of being a "most bullied" student is significantly lower in private schools than in public schools. But the estimated bullying risks are nonetheless non-trivial for private schools. Second, the largest local economies, i.e., mega-Manila cities and provinces, Cebu, and Davao City, appear to have the lowest estimated bullying risks, which may be contrary to expectations for areas with more crowded schools, i.e., with higher class sizes, especially in the public sector. Third, there is wide range of disparity in estimated bullying risks across schools even within provinces that locational averages may mask. In Abra, for example, while the average "most bullied" bullying risk in public schools is close to 40%, the trimmed range is close to 30%-points, suggesting substantial heterogeneity of estimated bullying risks across public schools in the province. In its neighboring province Ifugao, on the other hand, the average bullying risk is slightly lower, but the trimmed range is much tighter at around 10%-points, suggesting that the estimated bullying risk across public schools within the province are much similar. Fourth, such disparity in bullying risks appears to increase with average bullying risk by province.

6. Policy implications

The analyses we present in this study show that a non-trivial proportion of students are at risk of bullying in schools. The risk is higher for already disadvantaged students, particularly those who have experienced hunger and those who have repeated grade levels. Being female and enrolling in private schools appear to have some protective effects against school bullying. These highlight potential avenues for targeted measures to minimize school bullying exposure or to identify potential victims for post-bullying psychosocial interventions.

It is interesting that we document limited roles for the presence of school principals or guidance counselors in explaining the observed school bullying risks among students. This is in contrast with discussion in the literature highlighting the importance of these school personnel in preventing school bullying and providing support to victims (Brown, et al., 2020; Li, et al., 2017; Bauman, 2008). It should be noted, however, that our analysis may be gravely limited as we only accounted for the presence of school principals or guidance counselors in schools. It may be important to also document actual tasks and workload faced by these school personnel. Incorporating these in the analysis could further shed light on the mechanics and effectiveness of bullying case management in schools. Are they truly ineffective in preventing bullying and to foster conducive learning environments in schools? Does guidance personnels' case load affect the experience of bullying? Is there an ideal case load for guidance staff in schools? We leave these questions for further probing to future investigators.

The provincial small area estimates of "most bullied" bullying risk may be used as an initial guide to identify locations that require immediate attention. Provinces with high average but low dispersion of bullying risk among schools should be areas of concern as students are similarly at elevated risks of being school bullying victims. In any case, these areas of concern could be fertile grounds to pilot and learn from anti-bullying interventions that could eventually be rolled out to other schools.

School bullying is associated with learning losses in the Philippines. While the magnitude in foregone learning achievement appears to be minimal at around 0.05 standard deviations, their long-term economic implications could be quite substantial at PhP10-20 billion per year in lost

economic activity. This only underscores the importance of addressing school bullying, although this may be easier said than done. Unlike students' socio-economic status and other persistent household and community characteristics, school bullying experience may be a more readily modifiable risk factor of learning losses. Addressing bullying experience in schools, together with similarly modifiable risk factors, such as truancy and grade repetition, may be fruitful avenues for intervention to raise average learning achievement in the country.

It may be encouraging that the global evidence is replete with examples of anti-bullying interventions in schools that are found to be effective in lowering bullying perpetration and victimization. However, the estimated effects of these interventions are rather heterogeneous and are highly context specific. Finding what works in the Philippines' case necessarily requires identifying promising interventions, documenting its implementation, and fine-tuning them to the local realities that students face in schools across the country.

References

- Bauman, S. (2008). The role of elementary school counselors in reducing school bullying. *The Elementary School Journal*, 108(5), 362-375.
- Bernardo, A.B.I., M.O. II Cordel, M.O. Calleja, M.O., J.M.M. Teves, S.A. Yap and U.C. Chua (2024). Profiling low-proficiency science students in the Philippines using machine learning. *Humanities and Social Science Communications*, 10, 192.
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *Journal* of Human Resources. 8(4): 436–455.
- Brimblecombe, N., S. Evans-Lacko, M. Knapp, D. King, R. Takizawa, B. Maughan, and L. Arseneault (2018). Long term economic impact associated with childhood bullying victimization. *Social Science & Medicine*, 208, 134-141.
- Brown, J., J. Keesler, I. Karikari, G. Ashrifi, and M. Kausch (2022). School Principals Putting Bullying Policy to Practice. *Journal of Interpersonal Violence*, *37*(1-2), NP281-NP305.
- Carbone-Lopez, K., F.-A. Esbensen, and B.T. Brick (2010). Correlates and Consequences of Peer Victimization: Gender Differences in Direct and Indirect Forms of Bullying. *Youth Violence and Juvenile Justice*, 8(4), 332–350.
- Chiu, H., and E.J. Vargo (2022). Bullying and other risk factors related to adolescent suicidal behaviours in the Philippines: a look into the 2011 GSHS Survey. *BMC Psychiatry*, 22, 445.
- Darmawan, G.N., and A.A.S.S.K. Dharmapatni (2024). A multilevel analysis of student and school characteristics associated with 15-year-olds' reading performances: a Southeast Asian perspective. *Large-scale Assessments in Education*, 12, 40.
- Davies, S., G. Engström, T. Pettersson, and M. Öberg (2024). Organized violence 1989-2023, and the prevalence of organized crime groups. *Journal of Peace Research*, 61(4), 673-693.

- Drydakis, N. (2014), Bullying at school and labour market outcomes. *International Journal of Manpower*, 35(8), 1185-1211.
- Elvidge, C., K. Baugh, M. Zhizhin, C.H. Feng, and T. Ghosh (2017). VIIRS night-time lights. *International Journal of Remote Sensing*, 38(21), 5860-5879.
- Food and Nutrition Research Institute (2024). National Nutrition Survey, 2018-2021. Survey data. Taguig City, Philippines: DOST-FNRI.
- Gini, G., and T. Pozzoli (2009). Association Between Bullying and Psychosomatic Problems: A Meta-analysis. *Pediatrics*, 123(3), 1059–1065.
- Gini, G., T. Pozzoli, M. Lenzi, and A. Vieno (2014). Bullying Victimization at School and Headache: A Meta-Analysis of Observational Studies. *Headache: The Journal of Head and Face Pain*, 54(6), 976–986.
- Golmaryami, F.N., P.J. Frick, S.A. Hemphill, R.E. Kahn, A.M. Crapanzano, and A.M. Terranova (2016). The Social, Behavioral, and Emotional Correlates of Bullying and Victimization in a School-Based Sample. *Journal of Abnormal Child Psychology*, 44, 381–391.
- Halliday, S., T., Gregory, A. Taylor, C. Digenis, and D. Turnbull, D. (2021). The Impact of Bullying Victimization in Early Adolescence on Subsequent Psychosocial and Academic Outcomes across the Adolescent Period: A Systematic Review. *Journal of School Violence*, 20(3), 351–373.
- Hanushek, E.A. and L. Woessmann (2010). *The high cost of low educational performance: The long-run impact of improving PISA outcomes*. Paris, France: OECD.
- Herrin, A.N., J.B. Borja, N. Mayol, F.M. Largo, J.L.J. Alegado, I.N. Bas, and M.R.M. Abrigo (2024). Home and school environment component: Sense of belongingness and bullying for the Second Congressional Commission on Education (EdCom 2). PIDS Discussion Paper Series No. 2024-42. Retrieved in 30 November 2024 from https://pidswebs.pids. gov.ph/CDN/document/pidsdps2442.pdf
- Hosmer, D.W., and S. Lemeshow (2000). Applied logistic regression, 2nd Ed. New York: Wiley.
- Hosozawa, M., D. Bann, E. Fink, E. Elsden, S. Baba, H. Iso, and P. Patalay (2021). Bullying victimisation in adolescence: prevalence and inequalities by gender, socioeconomic status and academic performance across 71 countries. *eClinical Medicine*, 41, 101142.
- Jakubowski, M., T. Gajderowicz, and H.A. Patrinos (2023). Global learning loss in student achievement: First estimates using comparable reading scores. *Economics Letters*, 232, 111313.
- Jeong, S., D.-H. Kwak, B. Moon, and C. San Miguel (2013). Predicting School Bullying Victimization: Focusing on Individual and School Environmental/Security Factors. *Journal of Criminology*, 2013, 1–13.

- L. Chen, Y. Chen, Y., H. Ran, H., Y. Che, D. Fang, Q. Li, Y. Shi, S. Liu, Y. He, G. Zheng, and Y. Xiao (2024). Social poverty indicators with school bullying victimization: evidence from the global school-based student health survey (GSHS). *BMC Public Health*, 24, 615.
- Li, Y., P.Y. Chen, F.-L. Chen, and Y.-L. Chen (2017). Preventing school bullying: Investigation of the link between anti-bullying strategies, prevention ownership, prevention climate, and prevention leadership. *Applied Psychology*, 66(4), 577-598.
- Lian, Q., C. Yu, X. Tu, M. Deng, T. Wang, Q. Su, and X. Zuo (2021). Grade repetition and bullying victimization in adolescents: A global cross-sectional study of the Program for International Student Assessment (PISA) data from 2018. *PLoS Med*icine, 18(11): e1003846.
- Moore, S.E., R.E. Norman, S. Suetani, H.J. Thomas, P.D. Sly, J.G. Scott (2017). Consequences of bullying victimization in childhood and adolescence: A systematic review and metaanalysis. *World Journal of Psychiatry*, 7(1):60-76.
- Mukerjee, S. (2018). Childhood Bullying and Labor Market Outcomes in The United States. *Atlantic Economic Journal*, 46, 313–335.
- Nguyen, A.J., C. Bradshaw, L. Townsend, and J. Bass (2020). Prevalence and Correlates of Bullying Victimization in Four Low-Resource Countries. *Journal of Interpersonal Violence*, 35(19-20), 3767-3790.
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*. 14(3): 693–709.
- Odigie, Tracy, Esme Elsden, Mariko Hosozawa, Praveetha Patalay, and Jean-Baptiste Pingault. 2024. "The healthy context paradox: a cross-country analysis of the association between bullying victimisation and adolescent mental health." *European Child & Adolescent Psychiatry* 1-10.
- OECD (2023a). PISA 2022 Results (Volume II): Learning During and From Disruption. Paris: OECD Publishing.
- OECD (2023b). PISA 2022. Student database. Paris: OECD Publishing.
- OECD (2023c). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education, PISA*. Paris: OECD Publishing.
- OECD (2024). PISA 2022 Technical Report. PISA. Paris: OECD Publishing.
- Orbeta, A.C. Jr., K.A.M. Melad and M. Potestad (2020). Correlates of test performance of 15year-old students in the Philippines: Evidence from PISA. PIDS Discussion Paper Series No. 2020-57. Retrieved on 15 November 2024 from https://pidswebs.pids.gov.ph/ CDN/PUBLICATIONS/pidsdps2057.pdf
- Pengpid, S., and K. Peltzer (2019). Bullying victimization and externalizing and internalizing symptoms among in-school adolescents from five ASEAN countries. *Children and Youth Services Review*, 106, 104473.

- Sansone, R. A., J. S. Leung, and M.W. Wiederman (2012). Self-reported bullying in childhood: Relationships with employment in adulthood. *International Journal of Psychiatry in Clinical Practice*, 17(1), 64–68.
- Schoeler, T., L. Duncan, C.M. Cecil, G.B. Ploubidis, and J.-B. Pingault, J.-B. (2018). Quasiexperimental evidence on short- and long-term consequences of bullying victimization: A meta-analysis. *Psychological Bulletin*, 144(12), 1229–1246.
- Schonlau, M. (2005). Boosted regression (boosting): An introductory tutorial and a Stata plugin. *The Stata Journal*, 5(3), 330-354.
- van Geel, M., P. Vedder P, and J. Tanilon (2014). Bullying and Weapon Carrying: A Metaanalysis. *JAMA Pediatrics*, 168(8):714–720.
- Volk, A., W. Craig, W. Boyce, and M. King (2011). Adolescent risk correlates of bullying and different types of victimization. *International Journal of Adolescent Medicine and Health*, 18(4), 575-586.

Appendix A. Blinder-Oaxaca decomposition

We specify a linear education production function, which relates student's PISA score with individual, household, school and community inputs and other characteristics:

$$y_{aihsc} = b_0 + b_1 Z_i' + b_2 Z_h' + b_3 Z_s' + b_4 Z_c' + e_{aihsc},$$

where y_{aihsc} is the assessment score in test subject *a* for student *i* from household *h* living in community *c* and attending school *s*, while the vectors Z_i , Z_h , Z_s and Z_c are sets of characteristics corresponding to each indicated level. The variable e_{aihsc} is the model residual with standard linear regression properties. The coefficients b_k -s are conformable vectors that are estimated from data, which capture the relationship between test scores and a specific variable, conditional on other factors. The education production function above implies that the contribution of endowments to student achievement are additive and cumulates the contribution of individual, household, school and community factors.

Based on the linear regression above, we apply a Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973), wherein we decompose the difference in average scores of proficient and non-proficient students as follows:

$$E(y|P = 1) - E(y|P = 0)$$

$$= \underbrace{b \cdot [E(Z|P = 1) - E(Z|P = 0)]'}_{endowment \ effect} + \underbrace{E(Z) \cdot (b^{P=1} - b^{P=1})}_{coefficient \ effect} + \underbrace{[E(Z|P = 1) - E(Z|P = 0)] \cdot (b^{P=1} - b^{P=1})}_{interaction \ effect},$$

where P is an indicator variable indicating proficiency attainment by subject (proficient = 1, non-proficient = 0), and E(.) indicate the expectation operator. The estimated coefficients are stacked as b, with subscripts indicating the subsample on which it is estimated from. The endowments are similarly stacked as Z. The endowment effects provide a measure of the differences in average scores due group differences in endowments, while the coefficient effects measure the contribution of differences in estimated regression coefficients. Finally, the interaction effect account for simultaneous differences in endowments and regression coefficients. The coefficient and interaction effects together make up the unexplained effects.

The major domain assessed by PISA 2022 rotates on a three-cycle basis, where each of the core domains (Mathematics, Science, and Reading) takes its turn as the assessment's major domain (OECD, 2024). While the major domain for PISA 2022 is Mathematics, we also perform the Blinder-Oaxaca decomposition on Science and Reading scores to show differences in endowment and unexplained effects of various factors per core domain.

PISA 2022 defines Proficiency Level 2 as the baseline level that students need to fully participate in society (OECD, 2023c). Our analysis divides the sample in a similar manner: assigning students at Level 2 or higher as proficient and those at Level 1, non-proficient, for each assessment area (Mathematics, Science, and Reading).

Since assessment scores in PISA are imputed values from statistical models based on itemresponse theory (IRT) and that subsets of student background questions are asked randomly to students based on a within-construct matrix sampling (OECD, 2024), we performed the Oaxaca-Blinder decomposition using multiple imputation techniques that accounts for the random availability of the Z_k -s and the random draws of the IRT models.

In addition to multiply imputed data on PISA assessment scores and background questions, we also included school-level indicator variables derived from data available from DepEd's Learner Information System (LIS) and Basic Education Information System (BEIS) for SY 2021-2022. For schools with unavailable data for that school year, latest available LIS and/or BEIS data from previous school years were used.

Community-level variables were calculated from PSA's Census of Population and Housing (CPH) 2020. Additional community-level characteristics included in the model were total battle-related deaths from 2010 to 2019 taken from the Uppsala Conflict Data Program (Davies, et al., 2024), and the aggregated monthly records of nighttime lights in 2013 from the Earth Observation Group, collected using Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band (DNB) by Elvidge, et al. (2017).

Table A1.1 provides the full list and description of variables used in the decomposition. Means and standard deviations (S.D.) by student type, and standard mean differences (std. mean diff.) of outcomes are provided in Table A1.2. Finally, estimates of the Blinder-Oaxaca decomposition are provided in Table A1.3.

Variable	Description
ABGMATH	Ability grouping in mathematics classes, PISA 2022
BELONG	Sense of belonging. PISA 2022
BSMI	Expected occupation status, PISA 2022
BUILIED	Bullving exposure PISA 2022
COGACMCO	Encouraged mathematical thinking, PISA 2022
COGACINCO	Ensuring a mathematical reasoning, PISA 2022
	Population category 1 (< 3.000), PISA 2022
	Population category 2 ($3,000 - 15,000$), PISA 2022
COMMPOP3	Population category 3 (15.000 - 100.000), PISA 2022
	Population category 4 (100.000 - 1.000.000), PISA 2022
COMMPOP5	Population category 5 (1.000.000 - 10.000.000), PISA 2022
COMMPOPE	Population category 6 ($>$ 10 000 000), PISA 2022
	High curiosity PISA 2022
	Strict digital device policies. PISA 2022
DIGPREP	Prenaredness for digital learning PISA 2022
	Disciplinary climate PISA 2022
DURECEC	Duration of early childhood education. PISA 2022
EDUSHORT	Shortage of educational materials, PISA 2022
ENCOURPG	Encourages parent/guardian involvement. PISA 2022
FSCS	Socioeconomic status. PISA 2022
EXERPRAC	Practiced sports. PISA 2022
EXPECEDU	Expected level of education, PISA 2022
EXPO21ST	Exposure to 21st century math tasks, PISA 2022
EXPOFA	Exposure to formal & applied math tasks, PISA 2022
FAMSUP	Family support, PISA 2022
FAMSUPSL	Family support for self-directed learning, PISA 2022
FEELLAH	Feelings about learning at home, PISA 2022
FEELSAFE	Feeling safe at home, PISA 2022
GENDER	Gender, PISA 2022
INFOSEEK	Seeking information about future careers, PISA 2022
INSTLEAD	Level of instructional leadership, PISA 2022
LEARRES	Learning resources used while school was closed, PISA 2022
MACTIV	Availability of math-related extra-curricular activities, PISA 2022
MATHEASE	Perception of mathematics as easier than other subjects, PISA 2022
MATHEF21	Self-efficacy on 21st century math tasks, PISA 2022
MATHEFF	Self-efficacy on formal & applied math tasks, PISA 2022
MATHEXC	Available math extension courses, PISA 2022
MATHMOT	Motivation to do well in math, PISA 2022
MATHPERS	Effort and perseverance in math, PISA 2022
MATHPREF	Preference of math over other core subjects, PISA 2022
MCLSIZE	Math classroom size, PISA 2022
MISSSC	Missed school for more than 3 months, PISA 2022
MTTRAIN	Mathematics teacher training, PISA 2022
NEGSCLIM	Negative school climate, PISA 2022
PROATCE	Proportion of fully-certified teachers, PISA 2022

Table A1.1. Variable list and description

Variable	Description
PROBSCRI	Lack capacity to provide remote instruction, PISA 2022
PROBSELF	Problem with self-directed learning, PISA 2022
PROPMATH	Proportion of math teachers at school, PISA 2022
PROPSUPP	Proportion of personnel for pedagogical support, PISA 2022
RATCMP1	Availability of computers, PISA 2022
RATCMP2	Computers connected to internet, PISA 2022
RATTAB	Availability of tablet devices, PISA 2022
REPEAT	Grade repetition, PISA 2022
SCHAUTO	School autonomy in decision-making, PISA 2022
SCHRISK	Risks on safety, PISA 2022
SCHSEL	School selectivity, PISA 2022
SCHSIZE	School population, PISA 2022
SCHSUST	School efforts to sustain learning, PISA 2022
SCPREPAP	Readiness for Remote, after Pandemic, PISA 2022
SCPREPBP	Readiness for Remote, before Pandemic, PISA 2022
SCSUPRT	School closure support from other sources, PISA 2022
SCSUPRTED	School closure support from education authorities, PISA 2022
SDLEFF	Self-directed learning self-efficacy, PISA 2022
SISCO	Clear idea about future job, PISA 2022
SKIPPING	Skipped classes, PISA 2022
SMRATIO	Student to math teacher ratio, PISA 2022
SRESPCUR	School responsibility for curriculum, PISA 2022
STAFFSHORT	Shortage of teaching staff, PISA 2022
STDTEST	Use of standardized tests, PISA 2022
STUBEHA	Student behavior affecting school climate, PISA 2022
STUDYHMW	Studying outside school hours, PISA 2022
TARDYSD	Late arrival to school, PISA 2022
TCHPART	Teacher participation in school management, PISA 2022
TDTEST	Use of teacher-made tests, PISA 2022
ТЕАСНВЕНА	Teacher behavior affecting school climate, PISA 2022
TEACHSUP	Support from math teacher, PISA 2022
TEAFDBK	Regular feedback to teachers, PISA 2022
WORKHOME	Working at home, PISA 2022
WORKPAY	Working for pay, PISA 2022
acadvar	Academic variation (SD), NAT Y4 SY 11-12
agevar	Age variation (SD), LIS Official Enrollment SY 21-22
asstprinc	Existence of Assistant/Principal, BEIS SY 21-22
collegedad	Father at least college graduate, PISA 2022
collegemom	Mother at least college graduate, PISA 2022
conflictdeath	No. of Deaths due to Conflict, 2010-2019
density	Population density, PSA 2023
depratio_old	Dependency Ratio (Old), CPH20
depratio_young	Dependency Ratio (Young), CPH20
guidance	Existence of Guidance Counselor/Coordinator, BEIS SY 21-22
hsratio	Ratio of HS graduates in Municipality, CPH20

 Table A1.1. Variable list and description (continued)

Variable	Description
ipedratio	Ratio of IPEd students, LIS SY 21-22
lang	English as L1, PISA 2022
lhi	Language Homogeneity Index, CPH20
munmaleratio	Ratio of Male 15-24 y/o in Municipality, CPH20
nightlights	Night Lights, 2013
oldstuds	Ratio of Old students, BEIS SY 21-22
ppppratio	Ratio of 4Ps students, LIS SY 21-22
principal	Existence of Principal, BEIS SY 21-22
repratio	Ratio of Repeaters, LIS SY 21-22
sae	Municipal and City Level Small Area Poverty Estimates, 2015
scratio	Student-classroom ratio, BEIS & LIS SY 21-22
stratio	Student-teacher ratio, BEIS & LIS SY 21-22
transratio	Ratio of Transferee students, LIS SY 21-22
urb	Urbanicity, CPH20

 Table A1.1. Variable list and description (continued)

	Student type				Std.
Variable	Profici	ent	Non-proficient		Mean
	Mean	S.D.	Mean	S.D.	Diff.
Student					
Bullying exposure, PISA 2022	-0.251	1.018	0.713	1.177	-0.876
Missed school for more than 3 months, PISA 2022	0.055	0.229	0.346	0.476	-0.778
Grade repetition, PISA 2022	0.028	0.164	0.293	0.455	-0.775
Seeking information about future careers, PISA 2022	-0.193	0.848	0.209	1.059	-0.419
Skipped classes, PISA 2022	0.322	0.468	0.398	0.490	-0.159
Motivation to do well in math, PISA 2022	0.035	0.183	0.057	0.231	-0.107
English as L1, PISA 2022	0.060	0.238	0.070	0.255	-0.040
Late arrival to school, PISA 2022	0.691	0.760	0.719	0.716	-0.037
Perception of mathematics as easier than other subjects, PISA 2022	0.062	0.242	0.062	0.241	0.001
Clear idea about future job, PISA 2022	0.950	0.219	0.948	0.223	0.009
Problem with self-directed learning, PISA 2022	0.632	0.706	0.599	0.808	0.043
Father at least college graduate, PISA 2022	0.388	0.488	0.347	0.476	0.085
Self-efficacy on 21st century math tasks, PISA 2022	0.027	0.775	-0.045	0.897	0.086
Preference of math over other core subjects, PISA 2022	0.110	0.313	0.080	0.271	0.105
Gender, PISA 2022	0.573	0.495	0.498	0.500	0.151
Studying outside school hours, PISA 2022	6.956	2.777	6.430	3.042	0.181
Mother at least college graduate, PISA 2022	0.455	0.498	0.349	0.477	0.217
Self-directed learning self-efficacy, PISA 2022	0.032	0.766	-0.171	0.791	0.260
Self-efficacy on formal & applied math tasks, PISA 2022	-0.613	0.847	-0.858	0.994	0.266
Sense of belonging, PISA 2022	-0.156	0.779	-0.417	0.694	0.355
Expected level of education, PISA 2022	7.149	1.979	5.995	2.697	0.488
Effort and perseverance in math, PISA 2022	0.321	0.916	-0.221	1.095	0.537
Expected occupation status, PISA 2022	71.821	15.395	61.950	17.761	0.594
High curiosity, PISA 2022	0.593	1.016	-0.064	0.907	0.682

Table A1.2. Mean, standard deviation and standardized mean difference

	Student type				Std.
Variable	Proficient		Non-proficient		Mean
	Mean	S.D.	Mean	S.D.	Diff.
School					
Risks on safety, PISA 2022	-0.183	0.872	0.542	1.221	-0.684
Shortage of educational materials, PISA 2022	-0.016	1.345	0.797	1.322	-0.610
Ratio of Old students, BEIS SY 21-22	-0.338	0.550	0.058	1.047	-0.473
Ratio of 4Ps students, LIS SY 21-22	-0.352	0.685	0.060	1.032	-0.471
Lack capacity to provide remote instruction, PISA 2022	0.554	1.208	1.058	1.102	-0.436
Proportion of fully-certified teachers, PISA 2022	0.905	0.152	0.952	0.115	-0.347
Shortage of teaching staff, PISA 2022	-0.170	1.100	0.137	1.230	-0.263
Encourages parent/guardian involvement, PISA 2022	1.166	0.952	1.383	1.052	-0.216
Regular feedback to teachers, PISA 2022	1.889	1.176	2.121	1.156	-0.199
Level of instructional leadership, PISA 2022	0.973	0.875	1.127	0.866	-0.176
Student behavior affecting school climate, PISA 2022	-0.106	1.232	0.092	1.266	-0.158
School closure support from education authorities, PISA 2022	1.667	0.508	1.744	0.472	-0.156
Proportion of math teachers at school, PISA 2022	0.152	0.082	0.164	0.116	-0.128
Ratio of IPEd students, LIS SY 21-22	-0.058	1.101	0.010	0.981	-0.065
Math classroom size, PISA 2022	42.208	8.561	42.756	9.091	-0.062
Negative school climate, PISA 2022	-0.467	1.031	-0.412	1.029	-0.054
Teacher behavior affecting school climate, PISA 2022	-0.390	1.300	-0.325	1.442	-0.047
Use of teacher-made tests, PISA 2022	1.009	0.796	1.025	0.804	-0.020
Use of standardized tests, PISA 2022	0.866	1.105	0.858	1.137	0.007
School closure support from other sources, PISA 2022	1.619	0.505	1.605	0.494	0.027
Ability grouping in mathematics classes, PISA 2022	2.511	0.644	2.489	0.653	0.035
Availability of tablet devices, PISA 2022	0.217	0.593	0.192	0.468	0.048
Readiness for Remote, before Pandemic, PISA 2022	0.185	1.257	0.078	1.195	0.087
School population, PISA 2022	2898.238	3330.719	2625.495	2876.000	0.088

Table A1.2. Mean, standard deviation and standardized mean difference (continued)

	Student type				Std.
Variable	Profici	ent	Non-proficient		Mean
	Mean	S.D.	Mean	S.D.	Diff.
School					
Available math extension courses, PISA 2022	2.987	0.115	2.968	0.175	0.125
Proportion of personnel for pedagogical support, PISA 2022	0.154	0.168	0.127	0.196	0.149
Ratio of Repeaters, LIS SY 21-22	-0.033	1.157	0.006	0.971	-0.037
Support from math teacher, PISA 2022	0.637	0.894	0.474	1.038	0.168
Ratio of Transferee students, LIS SY 21-22	0.403	1.163	-0.069	0.952	0.444
Availability of math-related extra-curricular activities, PISA 2022	2.850	1.600	2.480	1.659	0.227
Preparedness for digital learning, PISA 2022	1.175	1.172	0.885	1.229	0.241
Mathematics teacher training, PISA 2022	0.952	0.450	0.808	0.695	0.247
Strict digital device policies, PISA 2022	0.893	1.095	0.588	1.301	0.254
Fostered mathematical reasoning, PISA 2022	0.446	0.809	0.183	1.189	0.258
School selectivity, PISA 2022	2.701	0.635	2.518	0.767	0.260
Exposure to 21st century math tasks, PISA 2022	0.609	0.844	0.370	0.909	0.273
Academic variation (SD), NAT Y4 SY 11-12	0.246	0.899	-0.044	1.011	0.303
Teacher participation in school management, PISA 2022	-0.149	1.253	-0.500	1.036	0.305
Encouraged mathematical thinking, PISA 2022	0.845	0.887	0.536	1.057	0.317
School autonomy in decision-making, PISA 2022	-0.238	1.130	-0.594	0.972	0.338
Readiness for Remote, after Pandemic, PISA 2022	0.636	0.484	0.409	0.800	0.343
Availability of computers, PISA 2022	0.342	0.347	0.224	0.313	0.357
School efforts to sustain learning, PISA 2022	-0.088	0.748	-0.406	0.797	0.410
Exposure to formal & applied math tasks, PISA 2022	0.529	0.570	0.251	0.685	0.440
Disciplinary climate, PISA 2022	0.119	0.716	-0.277	0.916	0.481
Age variation (SD), LIS Official Enrollment SY 21-22	0.469	1.321	-0.080	0.910	0.485
School responsibility for curriculum, PISA 2022	2.207	1.940	1.364	1.498	0.486
Computers connected to internet, PISA 2022	0.808	0.328	0.552	0.429	0.672

Table A1.2. Mean, standard deviation and standardized mean difference (continued)

	Student type				Std.
Variable	Profici	ent	Non-proficient		Mean
	Mean	S.D.	Mean	S.D.	Diff.
Home					
Working for pay, PISA 2022	0.509	1.679	2.394	3.356	-0.710
Family support for self-directed learning, PISA 2022	0.154	0.875	0.337	0.820	-0.217
Practiced sports, PISA 2022	3.947	3.196	4.663	3.624	-0.210
Feelings about learning at home, PISA 2022	0.525	0.866	0.508	0.876	0.020
Duration of early childhood education, PISA 2022	1.802	0.955	1.739	1.413	0.052
Working at home, PISA 2022	6.968	3.264	6.415	3.527	0.163
Feeling safe at home, PISA 2022	-0.211	0.839	-0.396	0.887	0.214
Family support, PISA 2022	0.027	0.965	-0.311	1.068	0.333
Learning resources used while school was closed, PISA 2022	0.506	0.727	0.144	0.868	0.452
Socioeconomic status, PISA 2022	-0.606	1.005	-1.460	1.098	0.811
Community					
Municipal and City Level Small Area Poverty Estimates, 2015	-0.492	0.781	0.084	1.009	-0.638
Dependency Ratio (Young), CPH20	-0.460	0.885	0.079	0.997	-0.572
Dependency Ratio (Old), CPH20	-0.150	0.903	0.026	1.013	-0.184
Ratio of Male 15-24 y/o in Municipality, CPH20	-0.222	0.904	0.038	1.011	-0.272
No. of Deaths due to Conflict, 2010-2019	-0.150	0.714	0.026	1.039	-0.198
Population category 1 (< 3,000), PISA 2022	0.004	0.062	0.023	0.151	-0.169
Population category 2 (3,000 - 15,000), PISA 2022	0.061	0.240	0.099	0.298	-0.139
Population category 3 (15,000 - 100,000), PISA 2022	0.269	0.444	0.462	0.499	-0.409
Population category 4 (100,000 - 1,000,000), PISA 2022	0.552	0.497	0.343	0.475	0.430
Population category 5 (1,000,000 - 10,000,000), PISA 2022	0.114	0.317	0.073	0.260	0.140
Population category 6 (> 10,000,000), PISA 2022	-	-	-	-	-
Language Homogeneity Index, CPH20	0.144	0.874	-0.025	1.018	0.178
Night Lights, 2013	0.210	1.067	-0.036	0.984	0.239
Population density, PSA 2023	0.400	1.213	-0.069	0.942	0.432

Table A1.2. Mean, standard deviation and standardized mean difference (continued)

Table A1.3. Blinder-Oaxaca decomposition

····	Math				Science		F	Reading	
	Ε'	Х	Т	E	Х	т	E	Х	т
Decomposed Effect	40.3	81.5	121.9	50.2	90.6	140.8	62.4	91.7	154.1
Variable Effects									
Student									
Bullying exposure	3.3	-0.5	2.8	4.7	-0.5	4.2	6.3	-0.9	5.4
Missed school for more than 3 months	4.4	0.4	4.8	6.5	0.2	6.7	4.6	0.7	5.4
Grade repetition	3.6	0.0	3.6	2.6	0.0	2.6	5.3	0.1	5.4
Seeking information about future careers	0.9	-0.4	0.5	1.3	-0.5	0.7	1.9	-0.9	1.0
Skipped classes	0.0	0.8	0.8	-0.3	1.4	1.2	0.0	1.3	1.3
Motivation to do well in math	0.0	0.3	0.4	0.0	0.5	0.5	0.2	0.4	0.6
English as L1	0.0	0.9	0.9	0.0	0.8	0.8	0.1	1.3	1.4
Late arrival to school	-0.1	-2.3	-2.3	-0.1	-2.2	-2.3	-0.1	-4.4	-4.5
Perception of math as easier than other subjects	0.0	0.7	0.7	0.0	0.2	0.2	0.0	0.7	0.7
Clear idea about future job	0.0	-6.8	-6.8	0.0	-3.4	-3.4	0.0	-3.5	-3.5
Problem with self-directed learning	0.0	-0.2	-0.2	0.1	-1.2	-1.1	0.1	-0.3	-0.2
Father at least college graduate	-0.4	3.4	3.0	-0.3	3.6	3.3	-0.5	4.4	3.9
Self-efficacy on 21st century math tasks	-0.2	-0.1	-0.3	-0.3	-0.1	-0.4	-0.3	-0.1	-0.4
Preference of math over other core subjects	0.1	-0.6	-0.5	-0.1	-1.0	-1.0	0.0	-1.5	-1.5
Gender	-0.2	-4.4	-4.6	-0.3	-4.7	-5.0	0.8	-6.2	-5.4
Studying outside school hours	0.2	-2.5	-2.3	-0.2	1.5	1.3	0.1	-3.9	-3.7
Mother at least college graduate	-0.9	3.1	2.2	-0.9	0.9	-0.1	-0.7	3.6	2.9
Self-directed learning self-efficacy	-0.1	-0.1	-0.2	-0.2	-0.1	-0.3	-0.1	-0.1	-0.2
Self-efficacy on formal & applied math tasks	0.1	-8.0	-7.9	0.1	-8.2	-8.0	-0.4	-7.1	-7.5
Sense of belonging	1.0	0.6	1.6	0.4	1.3	1.7	1.3	0.9	2.1
Expected level of education	1.1	-6.7	-5.6	1.7	-6.6	-4.9	1.8	-7.1	-5.3
Effort and perseverance in math	1.7	-0.7	1.0	0.8	-1.5	-0.7	1.3	-0.6	0.7

	Math			Science			Reading		
-	Ε'	Х	Т	E	Х	Т	E	Х	Т
Student									
Expected occupation status	0.6	7.4	7.9	0.8	7.5	8.3	0.9	0.9	1.8
High curiosity	4.2	-1.3	2.9	5.1	-0.1	5.0	6.8	-2.2	4.6
School									
Risks on safety	1.2	0.3	1.5	2.7	0.5	3.2	2.5	0.1	2.6
Shortage of educational materials	-0.1	0.0	-0.1	0.5	-0.1	0.4	0.5	-0.1	0.4
Ratio of Old students	0.7	0.0	0.7	1.4	-0.1	1.4	1.0	-0.5	0.5
Ratio of 4Ps students	0.2	0.7	0.9	0.6	1.6	2.3	1.0	0.6	1.7
Lack capacity to provide remote instruction	0.4	0.2	0.7	-0.5	0.0	-0.6	0.5	0.7	1.2
Proportion of fully-certified teachers	-0.4	7.9	7.6	-0.8	18.9	18.1	-0.5	-2.8	-3.4
Shortage of teaching staff	-0.1	0.0	0.0	-0.2	1.1	0.9	0.1	0.8	0.9
Encourages parent/guardian involvement	0.4	1.3	1.7	0.3	1.8	2.0	0.2	-2.7	-2.5
Regular feedback to teachers	0.1	-4.3	-4.2	0.2	-5.8	-5.6	0.1	-4.5	-4.5
Level of instructional leadership	-0.3	1.0	0.7	-0.1	3.1	3.0	0.0	2.5	2.5
Student behavior affecting school climate	-0.1	-0.6	-0.8	-0.2	-0.4	-0.5	-0.3	-0.4	-0.7
School closure support from education authorities	0.1	-0.8	-0.7	0.0	-4.0	-4.0	0.0	-12.6	-12.6
Proportion of math teachers at school	0.0	0.6	0.6	0.0	1.7	1.6	0.1	-2.7	-2.6
Ratio of IPEd students	0.0	0.3	0.2	0.0	0.5	0.5	-0.1	0.6	0.6
Math classroom size	0.0	-6.1	-6.1	0.0	1.3	1.3	0.0	-2.0	-2.0
Negative school climate	0.0	0.7	0.7	0.1	0.1	0.1	0.0	0.1	0.1
Teacher behavior affecting school climate	0.1	0.7	0.8	0.0	0.0	0.1	0.2	0.3	0.5
Use of teacher-made tests	0.0	-0.5	-0.6	0.0	-0.7	-0.7	0.0	3.2	3.2
Use of standardized tests	-0.1	-0.8	-0.9	-0.1	-1.0	-1.1	-0.2	-1.3	-1.5
School closure support from other sources	0.0	3.1	3.0	-0.1	4.6	4.6	-0.1	14.3	14.2
Ability grouping in mathematics classes	0.0	4.7	4.7	0.0	12.3	12.3	0.0	-0.4	-0.5
Availability of tablet devices	0.0	0.4	0.4	0.0	0.8	0.8	0.0	1.1	1.0
Readiness for Remote, before Pandemic	0.0	-0.5	-0.4	0.1	-0.6	-0.5	-0.1	-0.2	-0.3

	Math		Science			Reading			
	Ε'	Х	т	E	Х	т	E	Х	Т
School									
School population	0.0	-1.7	-1.7	0.1	-1.8	-1.7	0.3	-1.6	-1.3
Available math extension courses	0.0	56.5	56.4	0.1	80.9	80.9	0.1	39.8	39.9
Proportion of personnel for pedagogical support	-0.1	1.0	0.9	0.0	2.5	2.5	0.1	2.2	2.3
Ratio of Repeaters	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1
Support from math teacher	0.1	-0.2	-0.1	0.2	-1.5	-1.3	0.0	-2.6	-2.6
Ratio of Transferee students	0.4	-2.3	-2.0	0.1	-1.9	-1.8	1.0	-3.4	-2.4
Math-related extra-curricular activities	0.0	2.2	2.3	0.0	2.1	2.1	0.4	7.6	8.0
Preparedness for digital learning	0.1	0.9	1.0	0.5	-0.9	-0.4	0.1	-0.7	-0.6
Mathematics teacher training	0.0	-3.2	-3.2	0.0	-2.4	-2.3	-0.2	-3.0	-3.2
Strict digital device policies	0.4	-0.7	-0.3	0.4	0.3	0.6	0.3	-1.1	-0.8
Fostered mathematical reasoning	-0.1	0.2	0.1	-0.3	0.7	0.5	-0.1	0.3	0.2
School selectivity	0.1	-1.9	-1.8	0.3	-2.9	-2.6	0.0	-2.9	-2.9
Exposure to 21st century math tasks	0.4	-3.0	-2.6	0.4	-2.8	-2.3	0.3	-3.2	-2.8
Academic variation	-0.1	0.4	0.3	-0.3	-0.4	-0.7	-0.4	0.3	-0.1
Teacher participation in school management	-0.3	0.2	-0.1	-0.1	0.0	0.0	0.3	0.4	0.7
Encouraged mathematical thinking	0.7	-2.0	-1.3	0.7	-2.8	-2.1	1.3	-0.9	0.4
School autonomy in decision-making	0.3	-0.4	-0.1	0.2	-1.3	-1.1	-0.1	-0.4	-0.6
Readiness for Remote, after Pandemic	0.4	-0.3	0.1	0.5	1.1	1.6	0.4	0.6	1.0
Availability of computers	0.3	-0.4	0.0	0.2	-1.1	-0.9	0.3	0.7	1.1
School efforts to sustain learning	0.5	0.1	0.6	0.3	0.2	0.5	0.8	-0.1	0.7
Exposure to formal & applied math tasks	0.4	-0.2	0.1	0.9	0.5	1.3	0.7	0.1	0.8
Disciplinary climate	0.9	0.1	0.9	0.8	0.2	1.0	0.8	0.0	0.8
Age variation	2.5	-0.7	1.8	4.5	-0.5	3.9	5.2	-0.7	4.5
School responsibility for curriculum	0.4	2.6	3.0	0.8	-4.2	-3.4	0.3	4.1	4.4

	· · · ·	Math			Science			Reading		
		Ε'	Х	т	E	Х	Т	E	X	т
School										
	Computers connected to internet	1.2	2.5	3.7	2.5	5.9	8.4	1.3	2.3	3.6
Home										
	Working for pay	2.2	0.0	2.2	2.3	0.4	2.7	3.4	0.3	3.8
	Family support for self-directed learning	0.3	0.0	0.3	0.4	0.0	0.4	0.6	-0.7	-0.1
	Practiced sports	0.2	-2.5	-2.3	0.4	-3.5	-3.1	0.5	-6.1	-5.6
	Feelings about learning at home	-0.1	-2.2	-2.3	0.0	-2.4	-2.4	-0.1	-3.1	-3.2
	Duration of early childhood education	-0.1	4.1	4.0	-0.1	3.8	3.8	-0.1	5.5	5.4
	Working at home	0.3	-3.7	-3.3	0.2	-8.0	-7.8	0.7	-5.9	-5.2
	Feeling safe at home	0.0	0.2	0.3	0.1	0.3	0.4	0.0	0.8	0.8
	Family support	0.6	0.0	0.6	1.1	0.0	1.1	0.8	0.0	0.8
	Learning resources used while school was closed	0.0	1.4	1.4	0.1	1.1	1.3	0.0	1.7	1.7
	Socioeconomic status	6.0	0.8	6.8	7.4	-1.4	5.9	9.0	-0.2	8.9
Comm	unity									
	Small Area Poverty Estimates	0.0	-3.4	-3.4	-2.9	-3.0	-5.9	-2.2	-2.0	-4.1
	Dependency Ratio (Young)	-0.3	4.2	3.9	1.7	2.0	3.7	-0.4	3.9	3.5
	Dependency Ratio (Old)	-0.2	0.8	0.6	-0.1	0.7	0.6	-0.1	0.9	0.8
	Ratio of Male 15-24 y/o in Municipality	-0.2	-0.1	-0.3	-0.1	0.5	0.4	0.0	0.3	0.3
	No. of Deaths due to Conflict	0.1	-0.5	-0.4	0.2	-1.3	-1.2	0.3	-0.9	-0.6
	Population category 1 (< 3,000)	-	-	-	-	-	-	-	-	-
	Population category 2 (3,000 - 15,000)	0.0	0.0	0.1	-0.2	2.9	2.7	0.1	1.5	1.6
	Population category 3 (15,000 - 100,000)	-0.4	9.6	9.2	-1.6	10.8	9.2	0.6	3.7	4.2
	Population category 4 (100,000 - 1,000,000)	0.6	17.6	18.1	1.6	19.0	20.6	-0.9	7.0	6.1
	Population category 5 (1,000,000 - 10,000,000)	-0.1	5.6	5.5	0.3	6.2	6.4	-0.4	4.1	3.7
	Population category 6 (> 10,000,000)	-	-	-	-	-	-	-	-	-

	Math			Science			Reading		
	Ε'	Х	Т	E	Х	Т	E	Х	Т
Community									
Language Homogeneity Index	0.1	-0.1	-0.1	0.1	-0.1	0.0	0.2	-0.5	-0.3
Night Lights	0.6	-0.8	-0.2	0.1	-0.9	-0.8	0.7	-1.3	-0.7
Population density	0.8	-1.8	-1.1	0.6	-0.5	0.0	1.8	-2.7	-0.9
Urbanicity	0.2	-1.0	-0.8	0.5	1.0	1.5	0.8	-2.0	-1.2

Appendix B. Small area estimation of bullying risk

We specify several classification models for a Filipino student being part of the 2022 PISA's "most bullied" students. Following OECD (2023a), we tag a student in the Philippine round of the 2022 PISA as being part of the "most bullied" if the student belongs in the top ten percentile of students around the world by bullying exposure index. The index is based on nine questions on the PISA student background questionnaire that asks about frequency of exposure to different bullying activities. The index is a predicted score from a generalized partial credit model, and normalized relative to the OECD baseline.

We used simple linear regression and three common machine learning techniques to classify students by "most bullied" status. The ML techniques we employed include lasso (least absolute shrinkage and selection operator) and elastic net regression that build on the simple linear regression model by performing variable selection and regularization that aims to improve prediction accuracy. We also tuned a gradient-boosted classifier (GBC), also based on logistic regression, wherein the final regression tree is built on updates from a series of regression trees that improve model fit (Schonlau, 2005). Unlike the first three classifiers that require functional forms to be specified a priori, the GBC learns these relationships from the data through splits in the regression tree.

We include in the classifiers a list of predictor variables that have been previously documented as correlated with bullying victimization (e.g. Nguyen, et al., 2020; Volk, et al., 2006; Golmaryami, et al., 2016). However, we limit the variables to ensure maximum coverage among Filipino students in the 2022 PISA, and availability of school-level analogues that shall be used for out-of-sample prediction for the small-area estimation. Individual-level variables from PISA 2022 are supplemented by school indicators derived from DepEd's school administrative records and by community-level indicators calculated from PSA's 2020 CPH. The full list of predictor variables is presented in Figure 5 in the main report.

We assess the predictive ability of the classifiers by plotting their receiver operating curve (ROC), then calculating the area under the ROC curve (AUC). The results shown in Figure A2.1 suggest that all the classifiers have acceptable levels of predictive ability, with all AUCs being greater than the 0.7 threshold suggested by Hosmer and Lemeshow (2000). We performed formal tests of differences in the estimated AUCs of the classifiers. The result suggests that the AUCs are statistically indistinguishable from each other, with overlapping 95% confidence bounds across classifiers. This is despite the GBC performing qualitatively better based on the AUC.

We also performed an additional goodness-of-fit test on school-level "most bullied" bullying risk by regressing actual rates calculated from PISA 2022 and predicted school-level risks based on the classifiers. The expectation is that the estimated model constant is indistinguishable from zero, while the slope on the predicted risk is equal to one. The regression model estimates summarized in Table A2.1 show that while the GBC has both higher coefficient of determination and lower root-mean square error, its predictions drift quite substantially from the actual rates directly estimated from data.





Source: Authors' estimates based on 2022 PISA data. Note: The values shown in parentheses refer to estimated area under the receiver operating curve. Logistic – simple logistic regression; LASSO – least absolute shrinkage and selection operator; GBC – gradient-boosted classifier.

	Logit	LASSO	Elastic net	GBC
β1	0.985***	1.046***	1.084***	1.308***
	(0.044)	(0.050)	(0.050)	(0.041)
β0	0.012	-0.011	-0.025	-0.115***
	(0.017)	(0.019)	(0.019)	(0.016)
p-value (β1=1, β0=0)	0.503	0.418	0.158	0.000
RMSE	0.081	0.086	0.083	0.061
R-squared	0.735	0.699	0.720	0.847
N	187	187	187	187

Table A2.1. Model fit: Actual v	. predicted bully	ying rates
---------------------------------	-------------------	------------

Source: Authors' estimates based on 2022 PISA data. Note: The regression estimates above are used to test model fit between actual bullying rates estimated directly from PISA 2022 data and model predictions as specified. The expectation is that the linear fit lies on the 45 degree line, i.e., β 1=1, and that it crosses the y-axis at zero, i.e., β 0=0. *, **, *** denote statistical significance at the 10-, 5- and 1-percent alpha levels, respectively. Values in parentheses are standard errors. RMSE – Root-mean squared error; N – number of observations.

Based on the properties discussed above, we used the simple logistic regression model to predict school-level risk of a student being part of the 2022 PISA's "most bullied" around the world. In the school-level prediction, we replaced individual-level indicator variables with its expected value. In particular, we replaced the female, grade repetition and transferee student indicator variables with their respective shares among enrolled in the DepEd schools database. The hunger indicator is replaced by province-level share derived from the Household Food Insecurity Scale included in the 2018-2021 National Nutrition Survey (Food and Nutrition Research Institute, 2024).

The estimates are aggregated to the province-level by averaging school-level estimates weighted by enrollment. Separate estimates are provided for public and private schools to show differences in bullying risks across education sectors. We also calculated trimmed range within provinces by taking the difference between the 95th and 5th percentiles of bullying risks across schools to show the dispersion of bullying risks within locations.

Appendix C. Robustness analysis

We performed a similar analysis to Appendix B, but in this case including additional school variables that may influence bullying risk: (a) having a school principal (instead of a heat teacher) as school head, (b) employing a guidance counselor, (c) student-teacher ratio and (d) student-classroom ratio. Since these variables are only available for public schools in the DepEd schools database, we exclude private schools in our analysis.

Similar to Appendix B, we show the ROC curve in Figure A3.1, and the regression of actual school-level bullying risks with predicted values in Table A3.1. The results are qualitatively similar to those presented in Appendix B, although the AUCs are marginally smaller. We also calculated influence statistics, which we summarize in Figure A3.2. Except for the public school indicator, which we excluded in this analysis, the most influential variables that explain bullying risk for this sample subset remains to be hunger experience, grade repetition, dependency ratio, and sex. The additional school variables introduced in this analysis have influence statistics that are quite small relative to the identified most influential variables.



Figure A3.1. Receiver operating characteristic curve

Source: Authors' estimates based on 2022 PISA data. Note: The values shown in parentheses refer to estimated area under the receiver operating curve. The classifiers only include students in public schools. Logistic – simple logistic regression; LASSO – least absolute shrinkage and selection operator; GBC – gradient-boosted classifier.

Logit	LASSO	Elastic net	GBC
0.984***	1.024***	1.137***	1.303***
(0.060)	(0.069)	(0.076)	(0.051)
0.012	-0.005	-0.049	-0.116***
(0.024)	(0.028)	(0.031)	(0.021)
0.654	0.733	0.163	0.000
0.079	0.085	0.084	0.057
0.654	0.603	0.609	0.820
146	146	146	146
	Logit 0.984*** (0.060) 0.012 (0.024) 0.654 0.079 0.654 146	Logit LASSO 0.984*** 1.024*** (0.060) (0.069) 0.012 -0.005 (0.024) (0.028) 0.654 0.733 0.079 0.085 0.654 0.603 146 146	Logit LASSO Elastic net 0.984*** 1.024*** 1.137*** (0.060) (0.069) (0.076) 0.012 -0.005 -0.049 (0.024) (0.028) (0.031) 0.654 0.733 0.163 0.079 0.085 0.084 0.654 146 146

Table A3.1. Model fit: Actual v. predicted bullying rates

Source: Authors' estimates based on 2022 PISA data. Note: The regression estimates above are used to test model fit between actual bullying rates estimated directly from PISA 2022 data and model predictions as specified. The expectation is that the linear fit lies on the 45 degree line, i.e., β 1=1, and that it crosses the y-axis at zero, i.e., β 0=0. The analysis only include public schools. *, **, *** denote statistical significance at the 10-, 5- and 1-percent alpha levels, respectively. Values in parentheses are standard errors. RMSE – Root-mean squared error; N – number of observations.

Figure A3.2. Covariate influence



Source: Authors' calculations based on 2022 PISA data. Note: The influence statistic captures an explanatory variable's share of the log-likelihood of a regression tree model in a GBC. A higher share signifies greater importance in explaining the model. The model only includes students in public schools.