I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18318

The Effects of School Bullying Victimization on Cognitive, School Engagement, and Friendship Outcomes

Atsushi Inoue
Ryuichi Tanaka

DECEMBER 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18318

The Effects of School Bullying Victimization on Cognitive, School Engagement, and Friendship Outcomes

Atsushi Inoue
Nippon Institute for Research Advancement

Ryuichi Tanaka
University of Tokyo and IZA

DECEMBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effects of School Bullying Victimization on Cognitive, School Engagement, and Friendship Outcomes

This study investigated the effects of bullying victimization on cognitive, school engagement, and friendship outcomes using panel data collected from elementary school students in a Japanese city. Employing a value-added model that controls for prior outcomes, our findings revealed that bullying victimization significantly impairs both cognitive and school engagement and weakens friendship formation. Furthermore, a high prevalence of bullying victimization within the classroom was found to negatively impact cognitive outcomes in subsequent years. These findings underscore the importance of effective school bullying prevention in fostering human and social capital among school-aged children.

JEL Classification: I21

Keywords: school bullying, academic performance, school engagement, friendship, Japan

Corresponding author:
Ryuichi Tanaka
Institute of Social Science
The University of Tokyo
7-3-1 Hongo
Bunkyo-ku
Tokyo 113-0033
Japan
E-mail: ryuichi.tanaka@iss.u-tokyo.ac.jp

1. Introduction

In most advanced economies, primary and lower-secondary education is compulsory and forms the foundation for human capital development. These stages equip students with essential skills that are critical for further academic progression and long-term success in the labor market. The quality of education at these levels plays a pivotal role in shaping students' academic outcomes and economic prospects, underscoring the importance of maintaining high educational standards.

Despite its importance of ensuring high-quality education, bullying in schools remains a widespread and persistent global issue. According to the 2019 Trends in International Mathematics and Science Study (TIMSS), 29% of fourth graders and 23% of eighth graders across 64 countries reported experiencing bullying at least once a month, with 8% and 6% experiencing bullying weekly, respectively (Mullis et al., 2020). While Japan reported the lowest frequency of bullying among TIMSS participants at both grade levels, domestic trends revealed a troubling rise in the number of recognized cases. In the 2023 academic year, a total of 732,568 bullying cases were reported across all school levels, equating to 57.9 cases per 1,000 students—the highest rate recorded since 1985 (Ministry of Education, Culture, Sports, Science and Technology [MEXT], 2024).

Bullying has well-documented adverse effects on the mental and physical health of both perpetrators and victims, as well as on their educational trajectories and labor market outcomes. Extensive psychology and education literature has shown that students who experience bullying are more likely to suffer from anxiety, depression, deteriorating physical health, suicidal ideation, and an increased use of tobacco and illicit drugs (Halliday et al., 2021; Moore et al., 2017; Schoeler et al., 2018). Moreover, school violence and delinquency generate negative spillover effects that adversely affect students beyond those who are directly involved (Ahn & Trogdon, 2017; Carrell & Hoekstra, 2010; Figlio, 2007). The significant social costs associated with school violence underscore the urgent need for effective preventive and intervention strategies.[1]

Previous economic research has highlighted the detrimental effects of bullying victimization on academic performance, educational attainment, labor market outcomes (Ammermueller, 2012; Brown & Taylor, 2008; Eriksen et al., 2014; Gorman et al., 2021; Ponzo, 2013), and noncognitive skills (Mori and Onozuka, 2024; Sarzosa, 2024; Sarzosa & Urzúa, 2021). However, the effect of bullying victimization on friendship formation in schools remains unexplored.

Friendship formation can be conceptualized as a form of social capital within an educational context (Coleman, 1988).[2] Recent research has increasingly emphasized its importance, demonstrating that the size and characteristics of peer relationships significantly influence academic

[1] For instance, Pereznieto et al. (2010) estimate that the economic burden of school violence in the United States amounts to approximately $7.9 billion per year.

[2] The concept of social capital has been extensively explored in the sociological and economic literature (Bourdieu, 1986; Coleman, 1988; Putnam, 1993). Coleman (1988) introduced the concept in the educational context, emphasizing its role in linking social networks to educational outcomes.

performance (Fletcher et al., 2020; Hill, 2015; Lavy & Sand, 2019) and educational attainment (Mora & Oreopoulos, 2011; Patacchini et al., 2017). Lavy and Sand (2019) argue that three mechanisms are salient: (i) joint production—peer collaboration that provides complementary learning inputs and raises the marginal product of effort; (ii) social pressure—norms within friendship groups that shape effort and classroom conduct, thereby lowering the marginal cost of effort; and (iii) mutual insurance/social learning—informal support (e.g., sharing notes or explanations) that attenuates idiosyncratic productivity shocks such as absences or temporary setbacks.

Moreover, deterioration of friendship networks may depress human-capital accumulation not only through academic channels but also via mental-health and well-being. In economics, Lavy and Sand (2019) exploit quasi-random classroom assignments and show that separation from pre-existing friends (friendship loss) worsens noncognitive behavioural outcomes and lowers social well-being. Similarly, researches in sociology and child psychiatry report that having few or no friends—or heightened loneliness—is associated with more depressive symptoms (Ueno, 2005; Qualter et al., 2010).

Building on this understanding, this study extends the analysis of the consequences of bullying beyond cognitive outcomes and school engagement—treated as one facet of the broader noncognitive construct— to include the formation of friendships among students. In doing so, we aimed to provide a more comprehensive perspective on the effects of bullying victimization.

Using a value-added model that controlled for initial outcomes measured at the start of each school year, we assessed the impact of bullying victimization across multiple dimensions of student outcomes. The results indicated that students with lower academic performance, and fewer friendships were more likely to experience bullying victimization. Bullying victimization negatively affects academic performance, school engagement, and friendship formation. Specifically, a 1-standard-deviation (SD) increase in bullying victimization was associated with a 0.03 to 0.05 SD decline in test score growth, a 0.11 SD decrease in school engagement, and a 0.23 SD reduction in friendship formation. Moreover, high classroom-level bullying victimization negatively affected students' cognitive outcomes the following year.

This study contributes to the literature on human capital formation by examining the effects of school bullying victimization on educational outcomes. Previous studies have established significant links between bullying and adverse academic and labor market outcomes across various contexts. Brown and Taylor (2008) found that school bullying in the UK led to lower educational attainment and reduced wages in adulthood. Using data from 11 European countries, Ammermueller (2012) reported that bullying significantly lowers educational attainment, which in turn negatively affects labor market earnings. Eriksen et al. (2014) addressed the endogeneity of bullying in Danish data and found that bullying between the ages of 10 and 12 years significantly reduced ninth-grade grade point average (GPA). Gorman et al. (2021) demonstrated that bullying during junior high school

in England negatively affected academic performance, mental health, employment prospects, and income. Using Italian data, Ponzo (2013) found that bullying significantly lowers academic performance, with stronger effects at age 13 than at age 9.

With a special focus on noncognitive skills, Sarzosa and Urzúa (2021) and Sarzosa (2024) examined the effects of bullying on the dynamics of cognitive and noncognitive skill formation in South Korea. They found that students with weaker noncognitive skills were more likely to become victims of bullying. Furthermore, bullying at age 15 increases the likelihood of smoking and depression, reduces life satisfaction by age 18, and lowers university enrollment rates. Sarzosa (2024) further showed that bullying victimization at the age of 15 hinders the development of both cognitive and noncognitive skills, leading to widening skill gaps over time. The study also found that students with distinctive characteristics were more likely to be victims of bullying than their peers. These studies have shown that bullying in primary and secondary education can have significant adverse effects on both short- and long-term outcomes. However, few empirical studies have investigated the effects of bullying on friendship formation. Our study extends the analysis of bullying consequences beyond cognitive outcomes and school engagement to include the formation of friendships among students to provide a more comprehensive perspective on the effects of bullying victimization.

This study contributes to the strand of literature on bullying in Japan, which has been extensively conducted across various social science disciplines.[3] However, to the best of our knowledge, studies that have quantitatively examined the effects of bullying victimization on human capital outcomes, such as cognitive and noncognitive outcomes or career decisions, are limited because of the limited availability of longitudinal data linking bullying experiences to student outcomes in the Japanese context. As an exception, Tanaka and Morozumi (2019) analyzed the effect of additional teacher allocations on reported bullying using school-level data but found no significant effects. Nakamuro (2017) examined the relationship between class size and reported bullying cases and found no significant association. Tanaka et al. (2019) report that municipalities that transitioned early to a new educational board system experienced an increase in the number of reported bullying cases. Our study utilized panel data from a Japanese municipality to examine how bullying victimization in elementary school affects cognitive outcomes and school engagement, as well as friendship formation.

The remainder of this paper is organized as follows. Section 2 provides the institutional background for this study. Section 3 describes the estimation model used in this study. Section 4

[3] These studies have primarily been descriptive or correlational. For example, Morita et al. (1999) analyze survey data from over 7,000 participants, providing a comprehensive description of bullying dynamics from the perspectives of students, teachers, and parents. Hojo (2023) and Sudo (2014) use TIMSS data to examine the determinants of bullying victimization, reporting that gender and academic performance are associated with the likelihood of being bullied. In the field of education, studies have examined how teacher–student relationships influence bullying dynamics (Akiba, 2004; Akiba, Shimizu, & Zhuang, 2010), while comparative research has explored bullying in Japan and the UK (Kanetsuna & Smith, 2002).

describes the data used in this study. Section 5 examines the determinants of bullying victimization. Section 6 presents our estimation results. Section 7 explores the heterogeneity of the effects. Section 8 offers a discussion. Finally, Section 9 concludes the study.

2. Institutional Background

This section provides the institutional context relevant to the analysis.

2.1. Japan's Compulsory Education System

In Japan, compulsory education spans nine years, comprising six years of elementary school and three years of junior high school. Children enter elementary school in April, following their sixth birthday. Students are typically assigned to public schools based on their residential areas and are entitled to free education during this period, with no tuition or mandatory fees.[4] Although families may opt to attend private schools, most students attend designated public schools.[5]

In elementary school, students are assigned to a homeroom class in which they take most of the subjects together. Class formation generally considers various factors including students' individual characteristics and attitudes. The data used in this study, covering the academic year 2014–2016, reflected a maximum class size of 40 students.

2.2. Bullying Incidence in Japanese Schools

Over the past decade, the number of recognized bullying cases in Japanese schools has increased significantly. According to Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT, 2024), the total number of reported bullying incidents increased from approximately 186,000 in the 2013 school year to an unprecedented high of approximately 733,000 in 2023. In per capita terms, the incidence reached 57.9 cases per 1,000 students in 2023, the highest level since the records began. This rising trend briefly reversed during the 2020 school closures caused by the COVID-19 pandemic but subsequently resumed its upward trajectory.

Substantial regional variations exist across Japan. Prefectures such as Yamagata, Fukui, and Hokkaido had the highest incidence rates, exceeding 100 cases per 1,000 students. Conversely, the Ehime and Nagasaki Prefectures reported the lowest rates, below 20 cases per 1,000 students. Moreover, designated metropolitan areas reported higher incidence rates (73.0 cases per 1,000 students) than the national average, suggesting more frequent recognition of bullying in urban regions.

---

[4] Municipal Boards of Education in Japan are responsible for assigning students to elementary and junior high schools. Typically, each school has a designated school area, and students are allocated based on their residential address. However, some municipalities have introduced a school choice system to accommodate regional needs. As of May 2022, 23% of Boards of Education overseeing multiple elementary schools and 20% overseeing multiple junior high schools had implemented a school choice system (MEXT, 2022).

[5] According to the 2023 MEXT School Basic Survey, 98.1% of elementary school students in Japan were enrolled in public schools.

Bullying incidents are most frequently reported at the elementary school level. In 2023, approximately 589,000 cases (approximately 80% of all incidents) were recognized in elementary schools, compared to approximately 123,000 in junior high schools, and roughly 18,000 in high schools. The per capita incidence was notably higher in elementary schools (96.5 cases per 1, 000 students) than in junior high (38.1) and high schools (5.5).[6]

2.3. Japan's Anti-Bullying Measures

In July 2012, a student's death by suicide in Otsu City, Shiga Prefecture, attributed to bullying, underscored the urgent need for comprehensive anti-bullying measures. In response, the Act for the Promotion of Measures to Prevent Bullying was enacted in June 2013 and took effect in September of the same year.[7] This law established a framework for systematic anti-bullying initiatives requiring both the national government and schools to develop basic bullying prevention policies. Schools are also mandated to establish specialized organizations to implement these measures and ensure consistent prevention and intervention efforts.

2.4. Overview of the Academic Assessments Used in This Study

To address regional educational needs, some municipalities have conducted their own local academic assessments in addition to the National Assessment of Academic Ability, which has been administered by MEXT since 2007. Although these local assessments have more limited coverage than national assessments, they often include a broader range of grade levels and enable longitudinal tracking of the same student cohorts, providing detailed insights into student progress and learning outcomes.

This study uses data from a municipal academic assessment conducted in a Japanese city. The dataset follows a cohort of students who were in fourth grade in 2014 through the end of sixth grade. Assessments were conducted annually during the first or second week of April for all students in the municipality. While the assessed subjects varied slightly by grade, Japanese and mathematics were tested each year from fourth to sixth grade. Science was introduced in the fifth and sixth grades, and social studies was included in the sixth grade.

2.5. Overview of the Questionnaire-Utilities survey

The Questionnaire-Utilities survey, hereafter referred to as QU, is widely employed in Japanese

---

[6] It should be noted that the observed increase and regional variations might reflect not only differences in actual bullying occurrences but also variations in schools' efforts to identify and report even minor bullying incidents. (Tanaka et al., 2019).

[7] The Act defines bullying as:

> *Acts exerting a psychological or physical influence on a child that are committed by another child who attends the same school or otherwise has a certain personal relationship with the victimized child (including acts committed via the internet), resulting in mental or physical pain for the victim.*

schools as a diagnostic tool for assessing students' interpersonal relationships, satisfaction with school life, and learning motivation. Originally developed by Kawamura and Tagami (1997), this survey quantitatively captures students' satisfaction with school life, interpersonal relationships within classrooms, motivation towards school activities, and experiences of bullying and social isolation.

The survey included items similar to those found in international assessments such as the TIMSS and Program for International Student Assessment (PISA), covering aspects of bullying experiences, social isolation, self-esteem, and learning motivation. These scales have been adopted in educational and psychological research in Japan (Iida et al., 2021; Nishimura et al., 2022; Saito et al., 2015; Yamasaki et al., 2017).

In the QU, bullying victimization was measured using six questions that assessed whether students had experienced unpleasant remarks, violence, ridicule, social exclusion, or difficulties in group activities. School engagement was measured with a three-item index capturing happiness when mastering previously difficult material, enjoyment of answering questions or expressing opinions in class, and efforts to obtain good grades and improve academically. Friendship formation was assessed using nine questions, including whether students felt that their classmates were kind to them, considered them part of the group, provided encouragement, and offered recognition.

In the municipality analyzed in this study, the QU was administered twice a year, once between late May and early June, and again between late November and early December, targeting all students. However, sixth-grade students participated only in the first survey. For the schedule of academic assessments and QU, see Table A1.

3. Econometric Framework

To examine the effect of bullying victimization on educational outcomes, we estimated a value-added model of the education production function by incorporating class-fixed effects as specified below:

$$Y_{ic'st+1} = \beta_0 + \beta_1 Bullied_{icst} + \beta_2 Girl_{icst} + f(Y_{icst}) + \lambda_c + \varepsilon_{ic'st+1} \quad (1)$$

where $Y_{ic'st+1}$ represents the outcomes of student *i* in class *c*' at school *s* at the beginning of academic year t+1. $Bullied_{icst}$ indicates the extent of bullying victimization during academic year t, when the student was in class *c* at school *s*. $Girl_{icst}$ denotes the gender of the student, and $f(Y_{icst})$ is a fifth-order polynomial of the prior outcomes. λ_c represents the class-fixed effect, and $\varepsilon_{ic'st+1}$ is the error term.

The coefficient of interest is β_1, captures the effect of bullying victimization in academic year t on the outcomes at the beginning of academic year t+1. To address potential correlations

between bullying victimization within the same class, we used cluster-robust standard errors at the class level.

The class-fixed effect λ_c controls for fixed characteristics associated with the academic year, grade level, teacher, and school. This approach can mitigate the bias arising from classroom-specific factors that may simultaneously influence both bullying victimization and student outcomes, such as the nonrandom assignment of teachers, variations in teacher quality, and the presence of bullies. Including the polynomial function of the initial educational outcomes, $f\ (Y_{icst})$, as a covariate helps control for unobserved educational and family inputs accumulated over time. By accounting for $f\ (Y_{icst})$, we estimated the effect of bullying victimization on overall educational growth over one year.

We used the value-added specification without student-fixed effects. Prior literature emphasizes the persistent nature of bullying victimization due to self-reinforcing mechanisms (Sarzosa, 2024), indicating that controlling for prior outcome variables is necessary. In addition, it generated a high autocorrelation in bullying victimization and other outcome variables within students, leaving little variation in within-student outcome variables after controlling for student-fixed effects. By adopting a value-added modeling approach, we explicitly incorporated prior educational outcomes. This mitigates the endogeneity arising from the feedback mechanism inherent in bullying victimization by controlling for unobserved, time-invariant individual characteristics (such as innate ability and family background) and preserves sufficient variability in our key explanatory variable, bullying victimization.

We also acknowledge an important caveat of the value-added specification. Conditioning on lagged outcomes can exacerbate bias through mean reversion (Daw and Hatfield, 2018; Roth et al., 2023). If some non-victims experienced a large negative shock in the prior period, matching on the lagged outcome may select such units, who then improve mechanically in the next period. This would overstate the adverse effect of bullying when lagged outcomes are included. To gauge this concern, we present side-by-side estimates with and without lagged outcomes and discuss the differences. In addition, to assess the scope for bias from unobservables, we implement coefficient-stability analyses following Oster (2019).

4. Data

This study utilized panel data that combine individual student records from a city-specific academic assessment and the QU conducted in the city with a population of approximately 120,000. The dataset follows students who were in fourth grade at public elementary schools from 2014 through their sixth grade in 2016.

As shown in Table A2, municipal statistics indicate that there were slightly more than 1,000 fourth-grade students in 2014. Students with missing data for the variables used in the analysis were

excluded. Nevertheless, the final sample covered more than 95% of the students in each grade. Table 1 defines the variables used in the analysis, and Table 2 presents the summary statistics.

The outcome variables included combined test scores for Japanese and mathematics; individual test scores for Japanese, mathematics, and science; school engagement and friendship outcomes. All test scores were standardized to have a mean of 0 and SD of 1 for each subject and grade level. Because science tests were not administered in fourth grade, the number of observations for science was limited to students who took the tests in fifth and sixth grades. An analysis of the test score distributions indicated that they generally followed a normal distribution (Figure 1).

We use school engagement as an outcome capturing one dimension of noncognitive skills. Our clarification of the conceptual basis of our school-engagement index is as follows. Educational psychology views school engagement as a multidimensional construct with behavioral, emotional, and cognitive components (Fredricks et al. 2004). Our school-engagement index is constructed from three student responses collected in the first QU in May or June, as listed in Table A1. The index captures three facets: (i) a positive attitude toward mastering previously difficult material; (ii) active participation in class; and (iii) effort and goal orientation toward improving grades. According to the framework in educational psychology, the first facet maps primarily to emotional engagement, the second to behavioral engagement, and the third to behavioral engagement with some overlap in cognitive engagement. Hence, our composite is best interpreted as a measure of school engagement. We focus on this engagement-based index because closely related measures predict meaningful outcomes in the economics literature. For example, Dee and West (2011) use student-reported items—"Subject not useful for my future," "Do not look forward to subject," and "Afraid to ask questions in subject class" —as indicators of school engagement, a facet of noncognitive skills, and show that smaller classes improve these engagement measures. Furthermore, they show that higher levels of noncognitive skills, as captured by these engagement measures, are associated with a greater likelihood of full-time employment and higher earnings in adulthood.[8]

We sum the three school engagement items to create a composite score ranging from 3 to 12. For the analysis, this score was standardized to have a mean of 0 and a SD of 1, with higher values indicating stronger engagement. As shown in Figure 2, the raw distribution school engagement revealed that scores between 10 and 11 were most frequently observed across all grades, suggesting that many students responded positively to these questions.

[8] Relatedly, De Paola and Skatova (2024) construct an academic-motivation measure from items asking why students try to do well at school and why they do their homework, capturing external, introjected, identified, and intrinsic regulation, and show that this motivation predicts later standardized test scores. Jackson et al. (2020) construct two survey-based indices: a social index based on interpersonal skills and school connectedness, and a work-hard index based on effort, grit, and engagement. They find that fewer arrests and higher rates of high-school graduation, four-year college entry, and college persistence in schools that raise these indices.

The friendship outcome measure was also derived from the first QU, based on students' responses to nine specific items (see Table A1 for details).[9] Students responded to these items using the same 4-point Likert scale, and their responses were summed to create a composite score ranging from 9 to 36. For the analysis, this composite score was standardized to have a mean of 0 and a SD of 1, with higher values indicating stronger friendship outcomes. As shown in Figure 3, the raw distribution of friendship outcome scores was left-skewed, with the mode across all grades clustering at approximately 30 points. This suggests that many students had strong friendships with each other.

The key explanatory variable in this study was the extent of bullying victimization. This measure was constructed by summing students' responses to six survey questions related to bullying and social isolation (see Table A1 for details). The scores ranged from 6 to 24 per survey round. Since the QU was conducted twice a year, scores from both rounds were aggregated to capture annual bullying victimization, yielding a variable ranging from 12 to 48. To ensure robustness, the models were estimated separately using the first- and second-round scores. The bullying victimization measure was constructed at two time points: first in fourth grade, to estimate its effect on outcomes at the beginning of fifth grade, and again in fifth grade, to estimate its effect on outcomes at the beginning of sixth grade. In all estimations, the bullying victimization score was standardized to have a mean of 0 and a SD of 1 for each year, with higher scores indicating more severe bullying experiences. An analysis of the raw, unstandardized distribution of the annual bullying victimization scores revealed that the most common value was a minimum score of 12, with progressively fewer students reporting higher levels of bullying (see Figure 4).

Prior research indicates that children who experience bullying are likely to experience declines in educational outcomes, subsequently increasing their vulnerability to further victimization (Sarzosa, 2024). In this study, we calculated the autocorrelation coefficient of bullying victimization between fourth and fifth grades in our dataset and found it to be 0.642, indicating a substantial degree of persistence. Additionally, we examined the changes in raw bullying victimization scores (ranging from 6 to 24 points) between the two grades. Specifically, 16% of the students exhibited no change at all, 33% changed by one point or less, 46% by two points or less, and 56% by 3 points or less. More than half of the students experienced minimal or no variation in their bullying victimization scores. These results suggest a feedback mechanism in bullying victimization, implying that the severity of bullying remains relatively stable over time and leads to limited within-individual temporal variation.

A single data limitation applies to all three measures, namely bullying victimization, school engagement, and friendship. When the municipality provided the data, composite measures were available for nearly all students, whereas item-level responses were available for only about half of

[9] The nine items used here combine two separate subscales originally distinguished in the QU: the Peer Acceptance scale (six items) and the Friendship Relations scale (three items). Since both subscales measure aspects of peer relationships, we aggregated them into a single composite score for this analysis. Robustness checks confirm that using the two subscales separately produces qualitatively similar results.

the sample. We therefore use the composites in the main analysis. In the subsample with item-level responses, we reconstruct alternative measures and show that our main results are robust to the measurement choice.

In addition to individual-level data, class-level information, including class size and the proportion of female students, was incorporated to analyze the determinants of bullying victimization and the heterogeneity of its impacts.

5. Determinants of Bullying Victimization

This section examines the potential determinants of bullying victimization by analyzing both individual- and class-level characteristics. To identify individual-level factors, we regressed the bullying victimization scores from the second round of the QU on students' initial academic performance, school engagement, friendship, and gender. Using the second-round bullying victimization score mitigated simultaneity bias, as bullying victimization, school engagement, and friendship outcomes were measured at the same time. Since the second round of the QU was not administered to sixth-grade students, this analysis was limited to fourth- and fifth-grade students. To account for class-level heterogeneity, we included class-fixed effects in the regressions.

Table 3 presents the results. The findings indicated that students with higher academic performance were less likely to experience bullying (Table 3, Columns [1] and [4]). Similarly, students with higher school engagement in the first round of the QU were less likely to experience bullying in subsequent periods (Table 3, Column [2]). However, after controlling for academic performance and friendship outcomes, the relationship between school engagement and bullying victimization became positive. This result is likely driven by an omitted variable bias, as school engagement are correlated with both academic performance and friendship outcomes, which are associated with bullying victimization.[10] Friendship outcomes were associated with a lower likelihood of experiencing bullying (Table 3, Column [3]), and this result remained robust even after accounting for academic performance and school engagement (Table 3, Column [4]). Furthermore, as shown in Column (4) of Table 3, after controlling for academic performance, school engagement, and friendship outcomes, girls were more likely to experience bullying than boys.

Next, we examined class-level characteristics potentially associated with bullying victimization. Specifically, we regressed second-round bullying victimization scores on class-level characteristics, including class size and the proportion of female students. To identify these class-level associations, class- fixed effects were excluded from the regression. The results presented in Column (5) of Table 3 indicate that larger class sizes were associated with higher levels of bullying

[10] The correlation between school engagement and friendship outcomes is notably strong: while the correlation coefficient between academic performance and school engagement is 0.18, the correlation between school engagement and friendship outcomes is substantially higher at 0.55.

victimization, while a higher proportion of female students was associated with lower levels of bullying victimization. However, this relationship was not statistically significant.

In summary, bullying victimization was significantly and negatively associated with students' baseline academic performance and friendship outcomes, whereas it was positively associated with school engagement and female gender. Since these factors potentially influence both bullying victimization and subsequent academic achievement, this study estimated the impact of bullying victimization on educational outcome growth by controlling for these determinants and incorporating class-fixed effects.

6. Results

6.1. The Effect of Bullying on Academic Performance

We estimated the model specified in Equation (1), and the results are summarized in Table 4. Column (1) presents the results for the combined test scores for Japanese and mathematics, Column (2) for the Japanese language test scores, Column (3) for mathematics, and Column (4) for science. All models include fifth-order polynomials of initial test scores in year t and class-fixed effects. Standard errors are clustered at the class level.

These results consistently indicate that bullying victimization significantly hinders academic progress. Specifically, a 1 SD increase in bullying victimization is associated with a 0.03 to 0.05 SD decline in test score growth across all subjects. Notably, the negative impact of bullying victimization on academic performance was most pronounced in Japanese.

6.2. The Effect of Bullying on School Engagement

Next, we examined the effect of bullying victimization on school engagement. Table 5 presents the estimation results, which indicate that bullying victimization negatively affects school engagement. Specifically, a 1 SD increase in bullying victimization was associated with a 0.11 SD decline in the growth of school engagement.

6.3. The Effect of Bullying on Friendship Outcomes

Finally, we investigated the influence of bullying victimization on friendship outcomes. Table 6 presents the results of this analysis. Consistent with earlier findings, bullying victimization negatively affected friendship outcomes. Specifically, a 1 SD increase in bullying victimization was associated with a 0.23 SD reduction in friendship outcome growth.

6.4. Alternative specifications

One distinctive feature of our data is that students complete the QU twice each school year—once in May/June and again in November/December (Table A1). We exploit this structure by using

the first within-year QU report on bullying victimization as a baseline control and defining exposure as subsequent changes in reported victimization.

Let $\text{Bullied}_{t,1}$ denote the first (May/June) QU response in school year *t*, $\text{Bullied}_{t,2}$ the second (November/December) response in year *t*, and $\text{Bullied}_{t+1,1}$ the first QU response in year *t+1*. For academic outcomes—standardized test scores measured in April of *t+1*, we relate achievement to the within-year change in victimization during year *t*, $\Delta\text{Bullid}_t = \text{Bullied}_{t,2} - \text{Bullied}_{t,1}$, and include $\text{Bullied}_{t,1}$ as an additional baseline control. For school engagement and friendship—measured with the QU in May/June of *t+1* —we use both the within-year change ΔBullid_t and the cross-year change $\Delta\text{Bullid}_{t\to t+1} = \text{Bullied}_{t+1,1} - \text{Bullied}_{t,1}$, again conditioning on $\text{Bullied}_{t,1}$. All models include the same controls and fixed effects as in the main specification.

Table A3 reports the estimates. When test scores are the outcome, the coefficient on ΔBullid_t is negative across subjects and statistically significant for Japanese. For school engagement and friendship, both ΔBullid_t and $\Delta\text{Bullid}_{t\to t+1}$ are negative and statistically significant. In all specifications, the coefficient on the baseline control $\text{Bullied}_{t,1}$ is also negative and significant.

Taken together, after conditioning on baseline differences in bullying victimization and exploiting both within-year and cross-year variation, adverse effects of changes in victimization are evident for Japanese achievement, school engagement, and friendship. Moreover, the beginning-of-year level of victimization is negatively associated with all outcomes.

6.5. Robustness

Controlling for Subsequent Classroom Environments and Teacher Quality

We conducted robustness checks to examine the stability of our main findings under alternative model specifications.

The outcomes measured at the beginning of academic year *t+1* may have been influenced by the classroom environment in year *t+1*. If bullying victimization experienced in year *t* correlated with subsequent classroom environments, our estimates could potentially be biased. While academic assessments are administered within the first two weeks of the new academic year (early April), thereby minimizing this potential bias, the QU is conducted later, from late May to June, and may therefore reflect influences from the year *t+1* classroom environment. To address this concern, we estimate additional models that control for class-fixed effects in year *t+1*.

Teacher quality may also simultaneously affect bullying victimization and subsequent educational outcomes. To account for this possibility, we estimate models that incorporate teacher-fixed effects. However, since only a small number of teachers in our dataset taught classes in both fourth and fifth grades (five out of 71 teachers), most teacher-fixed effects were effectively absorbed by the class-fixed effects.

Table A4 reports the results of the robustness checks. The estimates indicate that controlling

for year *t+1* class-fixed effects (Columns [2], [5], and [8]) and teacher-fixed effects (Columns [3], [6], and [9]) does not substantively alter the bullying victimization coefficients across all outcome measures—cognitive, school engagement, and friendship outcomes—compared with our main results (Columns [1], [4], and [7]).

Thus, these findings confirm the robustness of our key results to potential confounding influences arising from classroom environments and teacher quality.

Robustness to Bullying Victimization Measured in Early and Late Academic Year

In our main analysis, we used an aggregated measure of bullying victimization, combining scores from both the first and second rounds of the QU to capture annual bullying exposure. Here, we further examined the robustness of these results by separately analyzing the effects of bullying victimization using the early measure $\text{Bullied}_{t,1}$ (first round, May/June) and the late-year measure $\text{Bullied}_{t,2}$ (second round, November/December) separately, replacing the annual aggregate with each measure in turn. This exercise assesses whether the timing of measurement—early versus late in the school year—alters the estimated relationships.

As shown in Table A5, our main findings remained robust under these separate measures. However, one exception arises when the bullying victimization measure from the later year survey is used: the coefficients for mathematics and science scores in the following year decrease in magnitude and lose statistical significance (see Table A5, Columns [9] and [10]). Nevertheless, the direction of the coefficients remained unchanged.

In summary, bullying victimization consistently and significantly impaired academic performance, school engagement and friendship outcomes.

Robustness to alternative constructions of key variables

In our main analysis, we used a single-dimensional index of bullying victimization constructed by summing Likert-scale responses to six questions. This is mainly because of the data limitations. As noted in Section 4, the municipality provided us the item-level responses for only about half of the sample, although the composite scores were available for nearly all students. However, because the use of alternative measures of bullying victimization may alter the results, we conduct a robustness check restricting the sample to observations with the item-level data and reconstructing the indices via factor analysis. We then re-estimate the main models using these factor-based measures. In addition, because the same concern may apply to other outcome variables originally formed by simple sums of items, we replicate this procedure for school engagement and friendship outcomes, constructing alternative factor-based indices for each.

Before conducting factor analyses, we verify that restricting the sample to students with item-level responses does not alter the main estimates due to sample selection bias when indices are

constructed as simple sums of the six Likert-scale items. Table A6 reports estimates of Eq. (1) for this subsample. As expected, standard errors are larger because the sample size is smaller. The coefficients on bullying victimization are modestly attenuated for Japanese achievement, science achievement, and school engagement, but they remain negative across all outcomes, and statistical significance persists for Japanese achievement, mathematics achievement, school engagement, and friendship. Overall, the main results are robust to limiting the analysis to the observations with item-level data. We treat the results in Table A6 as the benchmark for analyses that rely on the limited subsample with item-level responses.

Next, using the item-level subsample, we conduct factor analyses. In the first stage, we conduct an exploratory factor analysis (EFA) to determine the number of latent dimensions underlying the six bullying victimization items. In the second stage, we fit a confirmatory factor analysis (CFA) model, impose an EFA-informed factor structure, and obtain CFA-based factor scores.

For the bullying-victimization items, we estimate EFA models with one-factor and two-factor solutions. The EFA uses the polychoric correlation matrix and maximum likelihood estimation with oblimin rotation. Following Attanasio et al. (2025), we treat factor loadings of at least 0.40 as salient. As reported in Table A7, under the one-factor solution all six items load strongly (0.62–0.88), and the eigenvalue exceeds one, suggesting a single latent factor that captures overall bullying victimization; we therefore retain the one-factor solution as a benchmark. Under the two-factor solution, Factor 1 loads strongly on Items 1–3 (0.75–0.89) and moderately on Item 6, and its eigenvalue exceeds one, whereas Factor 2 loads strongly on Items 4–5 and moderately on Item 6, and its eigenvalue is below one. We nonetheless retain the two-factor solution alongside the one-factor model. This pattern aligns with the distinction, documented in the child development literature, between direct bullying (physical or verbal, face to face) and indirect bullying (relational or social exclusion) (e.g., Card et al., 2008; Crick and Grotpeter, 1995; Vaillancourt et al., 2003), which motivates the two-factor interpretation. Because Item 6 loads more strongly on Factor 2, we assign it to the indirect-bullying factor in the subsequent CFA. Because an EFA with a three-factor structure produced a Heywood case —i.e., an improper solution with a negative variance or a standardized loading or factor correlation greater than one (Brown, 2015)—, we do not consider three-factor models further.

In the CFA, we estimate two models based on the EFA results: (i) a one-factor model in which all six items load on a single latent bullying factor, and (ii) a two-factor model in which Items 1–3 load on a direct-bullying factor and Items 4–6 load on an indirect-bullying factor. We compare model fit using the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker–Lewis Index (TLI). Table A8 summarizes the results: across all criteria, the two-factor model fits slightly better than the one-factor model. As a robustness check, we complement the composite index derived from the one-factor model with alternative factor scores from the two-factor

specification and verify that our main results are robust to this alternative measurement.

Table A9 reports estimates of Eq. (1) using the standardized bullying-victimization score derived from the one-factor CFA. The bullying coefficients are negative across all outcomes and are statistically significant for every outcome except science achievement. These results indicate that replacing the simple-sum index with a factor-based index leaves the main conclusions unchanged.

Table A10 reports estimates based on standardized factor scores from the two-factor CFA that distinguishes direct from indirect bullying. Because the two factor scores are highly correlated (r = 0.893), we include each factor separately to avoid multicollinearity and unstable estimates. The coefficient on direct bullying (Table A10, columns [1], [3], [5], [7], [9], and [11]) and the coefficient on indirect bullying (Table A10,columns [2], [4], [6], [8], [10], and [12]) are negative in all specifications and statistically significant for all outcomes except combined Japanese–mathematics achievement (column [1]) and science achievement (columns [7], [8]). Across outcomes, the indirect-bullying coefficients tend to be somewhat larger in magnitude than the direct-bullying coefficients, suggesting that indirect victimization may be more detrimental to educational outcomes.

Next, we apply the same factor-analytic procedure to construct alternative indices for the school engagement and friendship outcomes. For school engagement (three items), we estimate a one-factor EFA using the same settings as above. Table A11 shows that all three items load strongly on a single factor. For friendship, we likewise estimate a one-factor EFA using the same settings; Table A12 indicates a similar pattern, with a single latent factor capturing overall friendship formation. An alternative two-factor EFA for the friendship items yielded a Heywood case; we therefore do not consider two-factor models further.

In the confirmatory factor analyses, informed by the EFA results, we fit (i) a one-factor model for school engagement in which all three items load on a single latent factor and (ii) a one-factor model for friendship in which all nine items load on a single latent factor. For school engagement, the CFA model is saturated and therefore exactly reproduces the sample covariance matrix (CFI = 1.000, TLI = 1.000, RMSEA = 0.000). By construction, these fit indices are uninformative; we therefore do not interpret them. For friendship, the model shows acceptable fit by conventional benchmarks (CFI = 0.961, TLI = 0.949, RMSEA = 0.074), supporting a unidimensional structure.[11]

Table A13 reports estimates of Eq. (1) when we replace the simple-sum indices with standardized factor scores from one-factor CFAs of the school engagement and friendship items. The estimates are similar to those in Table A6. Thus, we conclude that the main results are robust to using factor-analytic measures in place of item sums for both school engagement and friendship.

Robustness to excluding friendship-proximal items from the victimization index

[11] As a reference point, Hu and Bentler (1999) suggest that CFI and TLI values close to 0.95 and an RMSEA close to 0.06 indicate relatively good fit between a hypothesized model and the observed data.

Two of the six items used to construct our bullying victimization index—(4) *"Do you find yourself alone during recess or other breaks?"* and (5) *"Do you end up being the last one left out when groups are formed in class?"*—arguably capture social isolation rather than direct bullying and may also reflect friendship-related aspects. To address potential construct overlap, we recompute the victimization index excluding items (4) and (5), using the remaining four items, and restandardize the composite. We then re-estimate the baseline value-added specifications.

Table A8 reports the results. Across outcomes, coefficients on the alternative index remain uniformly negative. Magnitudes are modestly attenuated in absolute value relative to the baseline estimates in Table A6 (same sample with item-level responses), and statistical significance is retained for Japanese achievement, school engagement, and friendship. Overall, the findings are robust to excluding the two social-isolation items, alleviating concerns that our main results are driven by those components of the index.

7. Heterogeneous Effects

Next, we examined how the effects of bullying victimization varied according to class characteristics.

Specifically, we estimated models that include interaction terms between the bullying victimization variable and dummy variables for class-level characteristics, namely, class size and the proportion of female students. The class-level dummy variables take the value of 1 if they are above the median, and 0 otherwise.

Table 7 presents the estimation results. First, regarding heterogeneity by class size, the negative effects of bullying victimization on mathematics and science performance appeared to be offset for students in larger classes compared to those in smaller classes (Table 7, Columns [5] and [7]). Next, when examining the heterogeneity of bullying victimization effects by proportion of female students, the negative effects of bullying victimization on friendship outcomes were mitigated for students in classes with a higher proportion of female students (Table 7, Column [12]).[12]

8. Discussion

Testing the validity of the econometric framework

In Section 3, we noted that conditioning on lagged outcomes can amplify bias via mean reversion (Daw and Hatfield, 2018; Roth et al., 2023). To assess this concern, we estimate otherwise identical specifications with and without lagged outcomes on the same sample. Table A15 reports the results. Across outcomes, the bullying-victimization coefficient in the value-added specification is

[12] Additionally, we included interaction terms between bullying victimization and individual student characteristics (e.g., gender, prior academic performance, school engagement and friendship outcomes). However, none of these interactions yielded statistically significant results.

smaller in absolute value than in the corresponding specification without lagged outcomes. This pattern is opposite to what mean reversion would predict (which would yield more negative coefficients when lagged outcomes are included because non-victims selected on a transitory negative shock rebound mechanically). Hence, even if some mean-reversion bias could remain, its magnitude appears limited and does not overturn the gains in plausibility from controlling for prior outcomes.

Coefficient-stability analysis

If unobserved student traits in the regression models are correlated with bullying victimization, the estimated victimization coefficient may be biased. We therefore assess how strong selection on unobservables would have to be, assuming it operates in the same direction as selection on observables, to attenuate our estimates. Following Oster (2019), we compare the bullying coefficient and R-squared across a parsimonious specification with only sex and class fixed effects and a fuller value-added specification that additionally includes the lagged dependent variable. We report the coefficient-stability parameter δ, defined as the ratio of selection on unobservables to selection on observables required to drive the bullying coefficient to zero. Oster (2019) suggests that $\delta > 1$ can be considered robust. In implementation, we follow Oster (2019) and set the maximum attainable R-squared to one point three times the R-squared from the full specification and compute δ.

Table A15 reports δ by outcome, together with the R-squared values for the short and full specifications and the corresponding coefficients. These values imply that selection on unobservables would have to be approximately 0.87 to 2.5 times as strong as selection on observables to eliminate the estimated bullying effect. Although δ exceeds one for language achievement, science achievement, and school engagement, it is slightly below one for overall achievement (0.913), math achievement (0.869), and friendship outcomes (0.900), indicating that these results may not be robust to potential omitted-variable bias based on Oster's criterion. With this caveat, we would conclude that the full specification, which controls for the lagged dependent variable—one of the strongest predictors of outcomes in the next period—, helps reduce concerns that omitted-variable bias drives our main findings.

The Effects of Peer-Bullying Victimization on Student Outcomes

Thus far, we have examined how direct experiences of bullying victimization affect subsequent cognitive, school engagement, and friendship outcomes. Prior research suggests that the presence of disruptive classmates negatively affects other students' outcomes (Ahn & Trogdon, 2017; Carrell et al., 2018; Figlio, 2007; Kristoffersen et al., 2015; Neidell & Waldfogel, 2010; Zhao & Zhao, 2021).

Building on these findings, we investigated how the extent of bullying victimization among classmates influenced an individual student's outcomes in the following year. Specifically, we

estimate an alternative specification that extends Model (1) by adding the class-average bullying victimization measure among classmates, excluding the individual's own score denoted by $PeerBullied_{-icst}$ and removing class-fixed effects. We retained the individual-level bullying measure $Bullied_{icst}$ to separately identify the effects of an individual student's own bullying victimization and the average bullying victimization experienced by their classmates. The estimated model is as follows:

$$Y_{ic'st+1} = \beta_0 + \beta_1 Bullied_{icst} + \beta_2 PeerBullied_{-icst} + \beta_3 Girl_{icst} + \gamma X_{-icst} + f(Y_{icst}) + \varepsilon_{ic'st+1} \quad (2)$$

where X_{-icst} represents a vector of class-level peer attributes, including prior cognitive, school engagement, and friendship outcomes, the proportion of girls, and class size. We estimated the model without school-fixed effects because bullying victimization primarily varies between schools rather than within schools.[13] Given that school-level variation constitutes a more significant portion of the total variance in peer-bullying exposure, incorporating school-fixed effects eliminates most of the variation necessary for identifying meaningful effects. By controlling for a comprehensive set of class-level peer attributes, we effectively accounted for potential confounding factors while preserving sufficient variation to reliably estimate the impact of peer-bullying victimization on individual student outcomes.

Table 8 presents the estimation results. We found that a higher incidence of bullying victimization among classmates was significantly associated with lower cognitive outcomes, specifically the combined Japanese and mathematics scores in the following year (Table 8, Columns [1]). Although the effects were not statistically significant for the other outcomes, the results generally suggest that greater exposure to bullying victimization among classmates tends to be negatively associated with students' own cognitive outcomes in the subsequent year.

These findings suggest that not only does one's own bullying victimization matter, but also that being surrounded by peers who have experienced bullying can deteriorate students' future cognitive outcomes.[14]

---

[13] To examine whether bullying victimization varies more substantially between schools or across classes within schools, we conducted the following analysis. First, we calculated the deviation of each class's mean bullying score from its respective school mean, measuring within-school, across-class variation. Second, we calculated the deviation of each school's mean bullying score from the overall mean, measuring between-school variation. The SD of bullying victimization was 0.224 for the within-school, across-class component, compared to 0.269 for the between-school component, indicating greater variation between schools.

[14] Motivated by the possibility that victimization is concentrated among a few classmates, we replace the peer variable from the within-class average excluding student *i* with the within-class maximum of classmates' victimization excluding *i*. Across achievement, school engagement, and friendship outcomes, the coefficient on this maximum-based peer measure is statistically insignificant.

Spillovers and SUTVA: Implications for interpretation

Peer spillovers from classmates' bullying victimization may violate the Stable Unit Treatment Value Assumption (SUTVA) and thereby bias the baseline estimates in Tables 4–6. If spillovers are negative and own victimization is positively correlated with the leave-one-out class mean of victimization, omitting peer exposure induces downward bias—that is, the coefficient on own victimization becomes more negative than the direct effect. Accordingly, our student-level estimates should be interpreted as upper bounds in absolute value on the direct effect of own victimization.

As a complement, we estimate regressions aggregated to the classroom level:

$$\bar{Y}_{ct+1} = \beta_0 + \beta_1 \overline{Bullied}_{ct} + \gamma \bar{X}_{ct} + f(\bar{Y}_{ct}) + \varepsilon_{ct+1} \quad (3)$$

where $\bar{Y}_{ct+1}$ is the class-average outcome in year *t+1* for the set of students who were classmates in year *t* (individual outcomes are standardized before averaging); $\overline{Bullied}_{ct}$ is the class-average of standardized bullying victimization in year *t*; $f(\bar{Y}_{ct})$ is a fifth-order polynomial in the baseline class-average outcome; and $\bar{X}_{ct}$ includes the share of girls and class size. Standard errors are clustered at the school level.

Table A16 reports the estimates. Class-average victimization is negatively related to all outcomes; coefficients are statistically significant for all outcomes except science achievement and school engagement (Table A16, columns [4] and [5]). By construction, the classroom-level effect combines direct and peer spillover channels. These classroom-level results reinforce the individual-level evidence, indicating that higher victimization within a class depresses next-year class-average outcomes.

Comparing the Impact of Bullying Victimization to Class Size Reduction

Our main results indicate that a 1 SD increase in bullying victimization is associated with declines of approximately 0.03 to 0.05 SDs in academic growth, 0.11 SD in school engagement, and 0.23 SD in friendship outcomes. To contextualize these effect sizes, we compared them with the impact of class-size reduction on academic achievement, drawing on previous studies using Japanese data.

Among studies reporting statistically significant effects from class-size reduction in Japanese elementary and junior high schools, reducing class size by one student is associated with academic improvements of approximately 0.006 to 0.018 SD, depending on subject and grade level (Akabayashi & Nakamura, 2014; Hojo, 2013; Hojo & Senoh, 2019; Tanaka, 2020). Given these estimates, preventing a 1 SD increase in bullying victimization corresponds to the educational benefit of reducing

the class size by approximately two to five students.[15]

Considering that bullying victimization substantially impairs school engagement and friendship formation, the overall benefits of effective bullying prevention may be even greater. These findings underscore the importance and potential cost effectiveness of targeted anti-bullying measures as valuable components of educational policies.

9. Conclusion

This study used panel data from elementary school students in a Japanese municipality to examine the effects of bullying victimization on academic performance and noncognitive and friendship outcomes. Although previous research has extensively explored the effects of bullying on academic and noncognitive outcomes, few studies have specifically investigated its impact on friendships. By incorporating this dimension, our study provided a more comprehensive understanding of the overall consequences of bullying. Additionally, to our knowledge, no prior empirical research in Japan has directly estimated the effects of bullying victimization on educational outcomes. Thus, our findings complement the existing literature.

First, our analysis of the potential determinants of bullying victimization revealed that students with lower academic performance and weaker friendship outcomes are more likely to experience bullying. Next, to estimate the effects of bullying victimization, we employed a value-added model that controlled for prior outcomes given the high autocorrelation of bullying victimization. The results indicate that greater exposure to bullying is associated with significant negative effects on academic performance and noncognitive and friendship outcomes. Moreover, high levels of bullying victimization in the classroom worsened cognitive outcomes in the following year.

One potential mechanism by which bullying victimization negatively affects academic performance is the disruption of friendships. Previous studies have shown that positive peer networks contribute to improved academic performance (Fletcher, Ross, & Zhang, 2020; Hill, 2015; Lavy & Sand, 2019). At the same time, it is important to acknowledge that academic performance, noncognitive skills, and friendship outcomes are likely to be interconnected and mutually reinforced. For example, noncognitive skills may support both the formation of positive peer relationships and the development of academic competencies. Academic success can enhance students' self-confidence and social integration, thereby fostering strong friendships. Given these complex interlinks, we refrained from asserting a single definitive causal pathway. Rather, our findings highlight that bullying

[15] To interpret what a 1 SD increase in bullying victimization represents, we refer back to the original measurement scale. Our bullying-victimization measure consists of responses to 12 items (measured twice per year), each rated on a 4-point Likert scale, resulting in total scores ranging from 12 to 48 points. The mean bullying score among fourth and fifth graders is approximately 19.2 points, with an SD of 6.85 points. Thus, a 1 SD increase (6.85 points) represents a shift from around the mean (19 points) up to about 26 points. Concretely, a student initially responding "strongly disagree" to five items and "disagree" to seven items would, after this increase, respond "disagree" to 10 items and "slightly agree" to two items.

victimization can disrupt the dynamic interplay among the cognitive, noncognitive, and social domains, thereby impairing students' overall educational outcomes.

Finally, our results align with findings from the economics literature, highlighting the negative effects of bullying victimization on human capital accumulation. Reducing bullying not only alleviates distress for victimized students, but also likely promotes both the human and social capital of school-aged children.

Acknowledgement

This study is conducted as a part of the Project “Microeconometric Analysis of Education Policy” undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The draft of this paper was presented at RIETI DP seminar for the paper. We are grateful for the valuable comments on an earlier version of this paper from Yuki Onozuka, Eric Weese, Ayako Kondo and the participants in the RIETI workshop, and Daiji Kawaguchi, Daigo Nakata, and Eiichi Tomiura at the DP Seminar. We also wish to thank the anonymous municipality for the provision of data. The authors acknowledge financial support for this project from the Japan Society for the Promotion of Science (Grant Numbers 20H05629 and 25H00538 for Ryuichi Tanaka). Any errors are our own.

Disclosure Statement

Inoue declares that he has no relevant or material financial interests that relate to the research described in this paper. Tanaka declares that his research is supported by the Grant-in-Aid for Scientific Research of the Japan Society for the Promotion of Science (grant numbers: 20H05629 and 25H00538). There is no other conflict of interest to be disclosed for all the authors.

Reference

Ahn, T., & Trogdon, J. G. (2017) Peer delinquency and student achievement in middle school. *Labour Economics*, 44, 192-217.

Akabayashi, H., & Nakamura, R. (2014). Can small class policy close the gap? An empirical analysis of class-size effects in Japan. *The Japanese Economic Review*, 65, 253–281.

Akiba, M. (2004). Nature and correlates of Ijime—Bullying in Japanese middle school. *International Journal of Educational Research*, 41(3), 216-236.

Akiba, M., Shimizu, K., & Zhuang, Y. L. (2010). Bullies, victims, and teachers in Japanese middle schools. *Comparative Education Review*, 54(3), 369-392.

Ammermueller, A. (2012). Violence in European schools: A widespread phenomenon that matters for educational production. *Labour Economics*, 19(6), 908-922.

Attanasio, O. P., De Paula, Á., & Toppeta, A. (2025). Intergenerational mobility in socio-emotional skills. *Journal of Public Economics*, *248*, 105423.

Bentler, P. M. (1990a). Comparative fit indexes in structural models. *Psychological bulletin*, *107*(2), 238.

Bentler, P. M. (1990b). Fit indexes, Lagrange multipliers, constraint changes and incomplete data in structural models. *Multivariate behavioral research*, *25*(2), 163-172.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.) *Handbook of Theory and Research for the Sociology of Education*, 241-258. New York, Greenwood.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological methods & research*, *21*(2), 230-258.

Brown, S., & Taylor, K. (2008). Bullying, education and earnings: evidence from the National Child Development Study. *Economics of Education Review*, 27(4), 387-401.

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child development*, *79*(5), 1185-1229.

Carrell, S. E., & Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, 2(1), 211-228.

Carrell, S. E., Hoekstra, M., & Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, *108*(11), 3377-3415.

Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child development*, *66*(3), 710-722.

Coleman, J. S. (1988). Social capital in the creation of human capital. *American journal of sociology*, 94, S95-S120.

Daw, J. R., & Hatfield, L. A. (2018). Matching and regression to the mean in difference‐in‐differences analysis. *Health services research*, *53*(6), 4138-4156.

De Paola, M., & Skatova, E. (2024). Non-cognitive skills and social isolation in late childhood: An investigation of their impact on school performance in Italy. *Economics of Education Review*, *103*, 102581.

Dee, T. S., & West, M. R. (2011). The non-cognitive returns to class size. *Educational Evaluation and Policy Analysis*, *33*(1), 23-46.

Eriksen, T. L. M., Nielsen, H. S., & Simonsen, M. (2014). Bullying in elementary school. *Journal of Human Resources*, 49(4), 839-871.

Figlio, D. N. (2007). Boys named Sue: Disruptive children and their peers. *Education finance and policy*, 2(4), 376-394.

Fletcher, J. M., Ross, S. L., & Zhang, Y. (2020). The consequences of friendships: Evidence on the effect of social relationships in school on academic achievement. *Journal of Urban Economics*, 116, 103241.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research*, *74*(1), 59-109.

Gorman, E., Harmon, C., Mendolia, S., Staneva, A., & Walker, I. (2021). Adolescent school bullying victimization and later life outcomes. *Oxford Bulletin of Economics and Statistics*, 83(4), 1048-1076.

Halliday, S., Gregory, T., Taylor, A., Digenis, C., & 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.

Hill, A. J. (2015). The girl next door: The effect of opposite gender friends on high school achievement. *American Economic Journal: Applied Economics*, 7(3), 147-177.

Hojo, M. (2013). Class-size effects in Japanese schools: A spline regression approach. *Economics Letters*, 120(3), 583–587.

—— (2023). *Shoninzu gakkyu no keizaigaku* (The Economics of Small Class Sizes). Tokyo: Keio University Press.

Hojo, M., & Senoh, W. (2019). Do the disadvantaged benefit more from small classes? Evidence from a large-scale survey in Japan. *Japan and the World Economy*, 52, 100965.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Iida, J., Ito, A., Aoyama, I., Sugimoto, K., Endo, H., Chan, M., & Furlong, M. J. (2021). Validation of a social emotional wellness survey among Japanese elementary school students. *The Educational and Developmental Psychologist*, *38*(1), 121-130.

Jackson, C. K., Porter, S. C., Easton, J. Q., Blanchard, A., & Kiguel, S. (2020). School effects on socioemotional development, school-based arrests, and educational attainment. *American Economic Review: Insights*, *2*(4), 491-508.

Kanetsuna, T., & Smith, P. K. (2002). Pupil insights into bullying, and coping with bullying: A bi-national study in Japan and England. *Journal of School Violence*, 1(3), 5-29.

Kawamura, S., & Tagami, F. (1997). Ijime higai/gakkyu futekio hakken shakudo no sakusei [Development of a scale to detect bullying and isolation in class]. *Japanese Journal of Counseling Science (Kaunseringu Kenkyu)*, 30, 112–120. (in Japanese)

Kristoffersen, J. H. G., Krægpøth, M. V., Nielsen, H. S., & Simonsen, M. (2015). Disruptive school peers and student outcomes. *Economics of Education Review*, *45*, 1-13.

Lavy, V., & Sand, E. (2019). The effect of social networks on students' academic and non-cognitive behavioural outcomes: evidence from conditional random assignment of friends in school. *The Economic Journal*, 129(617), 439-480.

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2024). *Reiwa 5 nendo Jidou Seito no Mondai Koudou Futoukou tou Seito Shidoujou no Shokadai ni Kansuru Chousa Kekka ni Tsuite* (Results of the Survey on Problematic Behavior, School Non-attendance, and Educational Challenges in 2022). (in Japanese).

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2023). *Reiwa 4 nendo Kouritsu Shou Chuugakkou tou ni okeru Kyouik Katei no Hensei Jisshi Joukyou Chousa no Kekka ni Tsuite* (Results of the Survey on the Organization and Implementation of Curricula in Public Elementary and Junior High Schools in 2022). (in Japanese).

Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2022). Shugakko no shitei, kuikigai shugaku no katsuyo jokyo chosa ni tsuite (Survey on the utilization of school designations and out-of-zone schooling) (in Japanese).

Moore, S. E., Norman, R. E., Suetani, S., Thomas, H. J., Sly, P. D., & Scott, J. G. (2017). Consequences of bullying victimization in childhood and adolescence: A systematic review and meta-analysis. *World journal of psychiatry*, 7(1), 60.

Mora, T., & Oreopoulos, P. (2011). Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, 30(4), 575-581.

Mori, H., & Onozuka, Y. (2024). One Size Fits All? The Adverse Effects of Strict School-Entry Policy on Noncognitive Development. Presented at the Japanese Economic Association 2024 Fall Meeting.

Morita, Y., Taki, M., Hata, M., Hoshino, K., & Wakai, S. (1999) *Nippon no Ijime* (Bullying in Japan) (in Japanese). Tokyo: Kaneko Shobo.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020) *TIMSS 2019 International Results in Mathematics and Science.* Retrieved from Boston College.

Nakamuro, M. (2017). Shoninzu gakkyu wa ijime, boryoku, futoko o herasu no ka (The Effect of Class

Size Reduction on Bullying, Violent Behavior, and Truancy). (in Japanese). RIETI Discussion Paper Series, 17-J-014.

Neidell, M., & Waldfogel, J. (2010). Cognitive and noncognitive peer effects in early education. *The Review of Economics and Statistics*, *92*(3), 562-576.

Nishimura, T., Murakami, T., & Sakurai, S. (2022). Interpersonal predictors of loneliness in Japanese children: variable-and person-centered approaches. *BMC psychology*, *10*(1), 164.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187-204.

Patacchini, E., Rainone, E., & Zenou, Y. (2017). Heterogeneous peer effects in education. *Journal of Economic Behavior & Organization*, *134*, 190-227.

Pereznieto, P., Harper, C., Clench, B., & Coarasa, J. (2010). *The Economic Impact of School Violence,* Plan International & Overseas Development Institute.

Ponzo, M. (2013). Does bullying reduce educational achievement? An evaluation using matching estimators. *Journal of Policy Modeling*, 35(6), 1057-1078.

Putnam, R. D. (1993). The prosperous community: Social Capital and Public Life. The American Prospect, 4(13), 35-42.

Qualter, P., Brown, S. L., Munn, P., & Rotenberg, K. J. (2010). Childhood loneliness as a predictor of adolescent depressive symptoms: an 8-year longitudinal study. *European child & adolescent psychiatry*, *19*(6), 493-501.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, *235*(2), 2218-2244.

Saito, E., Watanabe, M., Gillies, R., Someya, I., Nagashima, T., Sato, M., & Murase, M. (2015). School reform for positive behaviour support through collaborative learning: utilising lesson study for a learning community. *Cambridge journal of education*, *45*(4), 489-518.

Sarzosa, M. (2024). Victimization and skill accumulation: The case of school bullying. *Journal of Human Resources*, 59(1), 242-279.

Sarzosa, M., & Urzúa, S. (2021). Bullying among adolescents: The role of skills. *Quantitative economics*, 12(3), 945-980.

Schoeler, T., Duncan, L., Cecil, C. M., Ploubidis, G. B., & Pingault, J. B. (2018). Quasi-experimental evidence on short-and long-term consequences of bullying victimization: A meta-analysis. *Psychological bulletin*, 144(12), 1229.

Sudo, K. (2014). Ijime to gakuryoku: TIMSS 2011 chugakusei deta no keiryo bunseki kara (Bullying and Academic Performance: A Quantitative Analysis of TIMSS 2011 Junior High School Data) (in Japanese). *Bulletin of Edogawa University*, 24, 121–129.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, *25*(2), 173-180.

Tanaka, R. (2020). Konkyo o katsuyo shita kyoiku seisaku keisei e mukete: Jichitai kyoiku data o mochiita class-size shukusho koka no kensho [Toward the Evidence-Based Education Policy Making by Local Governments - An Analysis of Heterogenous Effects of Class-size Reduction on Students[1]Educational Achievements Using Administrative Panel Data]. *Journal of Social Security Research (Shakai Hosho Kenkyu)*, 5(3), 325–340. (in Japanese)

Tanaka, R., Bessho, S., & Morozumi, J. (2019). Shin kyoiku iinkai seido ga ijime no ninchi kensu ni ataeta eikyo ni tsuite: Tokyo-to no kushichoson-betsu data o mochiita bunseki [Accountability in Education Policy and Recognized Cases of School Bullying: An Analysis of Municipality-Level Data in Tokyo]. *Financial Review (Fainansharu Rebyu)*, 141, 44–64. (in Japanese)

Tanaka, R., & Morozumi, J. (2021). Kyoin kahai to mondai kodo: Katei keizai kankyo no yakuwari [Additional Teacher Allocation and Problematic Behavior of Students: A Role of Households' Economic Environment]. *JCER Economic Journal (Nihon Keizai Kenkyu)*, 79, 93–126. (in Japanese)

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1-10.

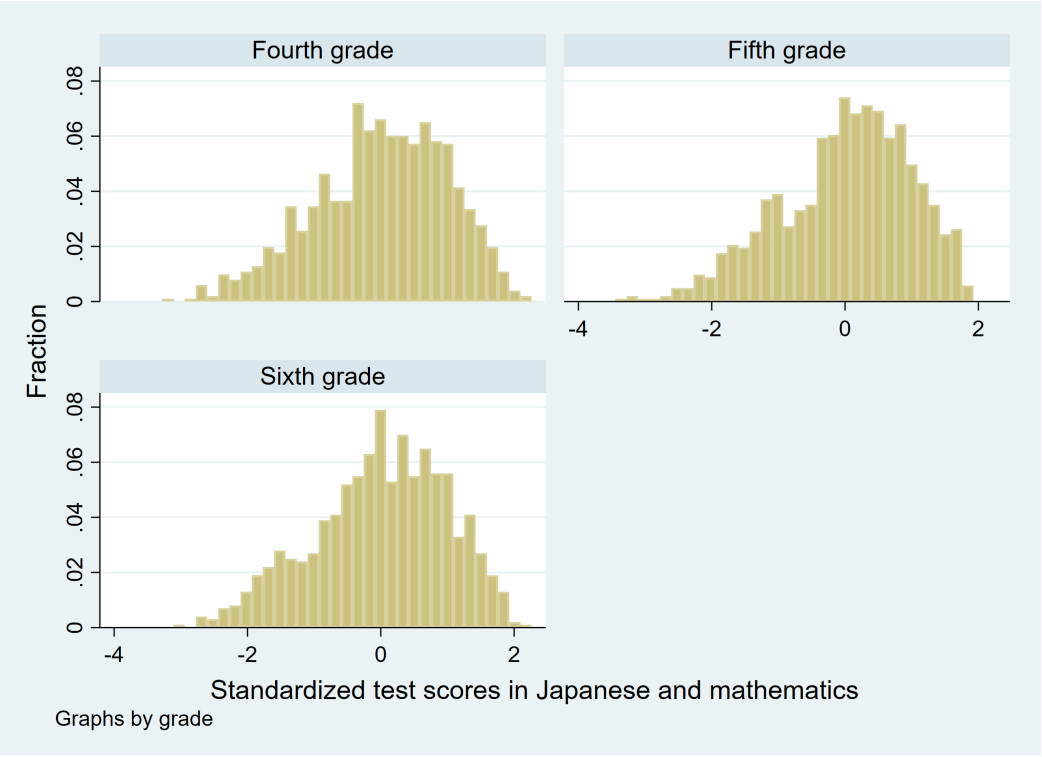
Ueno, K. (2005). The effects of friendship networks on adolescent depressive symptoms. *Social science research*, *34*(3), 484-510.

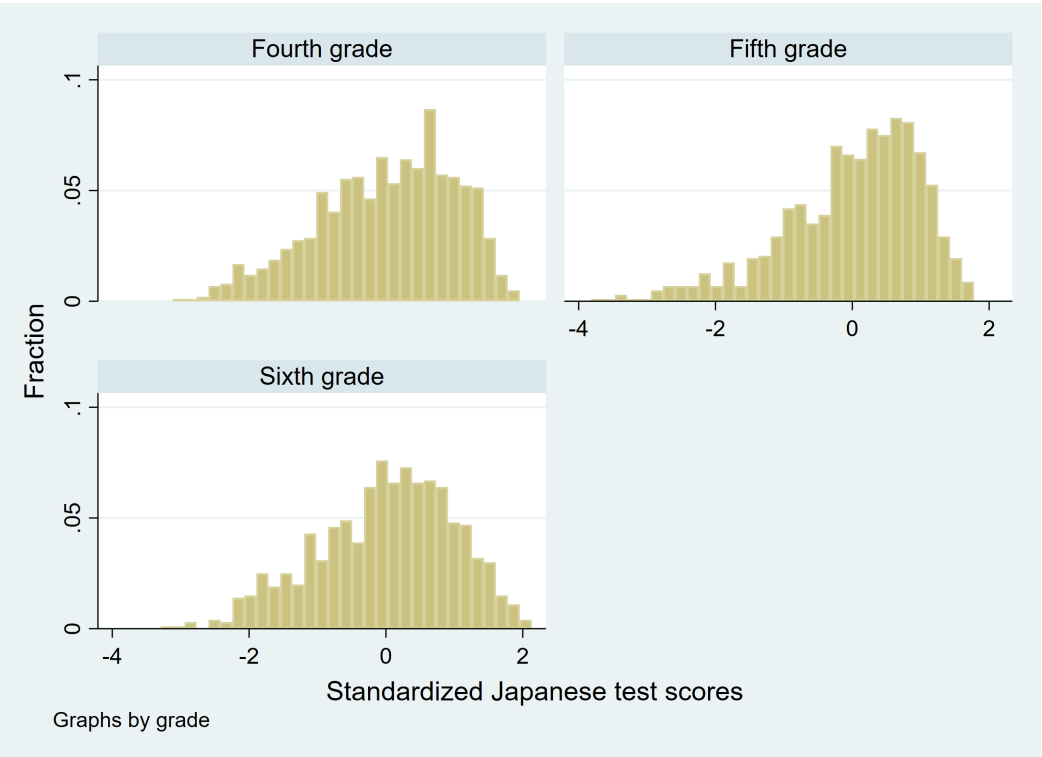
Vaillancourt, T., Brendgen, M., Boivin, M., & Tremblay, R. E. (2003). A longitudinal confirmatory factor analysis of indirect and physical aggression: Evidence of two factors over time?. *Child development*, *74*(6), 1628-1638.

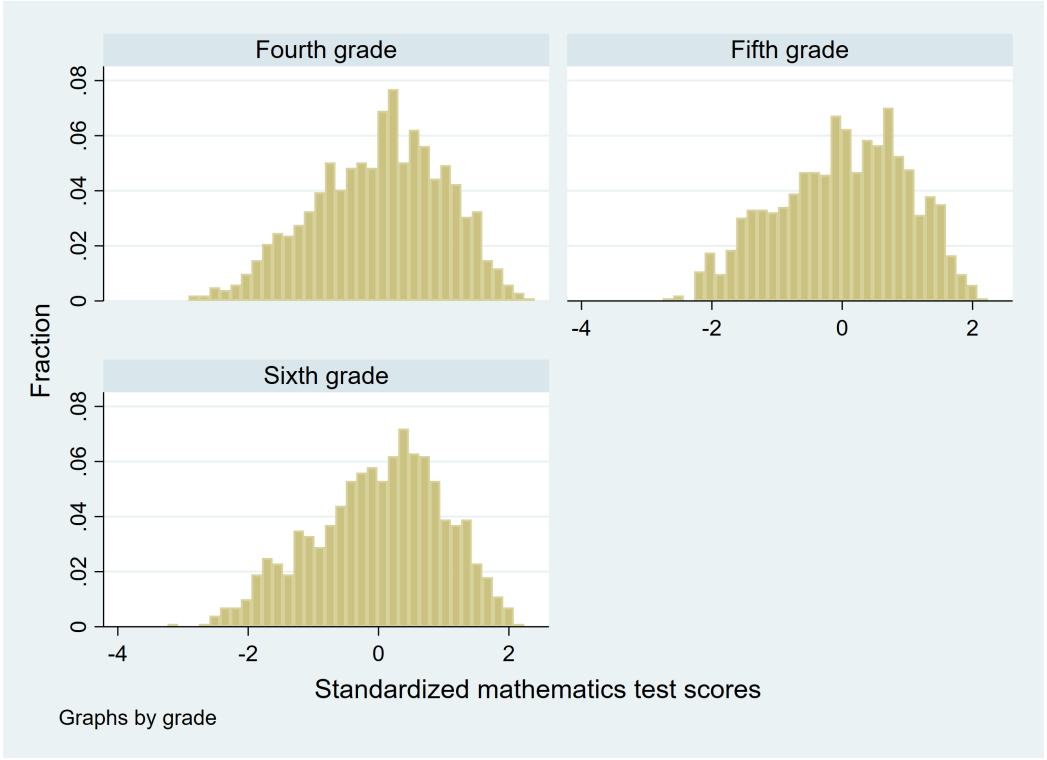
Yamasaki, K., Umakoshi, A., & Uchida, K. (2017). Efficacy of a school-based universal program for bullying prevention: considering the extended effects associated with the achievement of the direct purposes of the program. *Int'l J. Soc. Sci. Stud.*, *5*, 1.

Zhao, L., & Zhao, Z. (2021). Disruptive peers in the classroom and students' academic outcomes: Evidence and mechanisms. *Labour Economics*, 68, 101954.

Figure 1. Distribution of test scores.






Fourth grade
Fifth grade
Sixth grade
Fraction
0
.02
.04
.06
.08
-4
-2
0
2
Standardized mathematics test scores
Graphs by grade

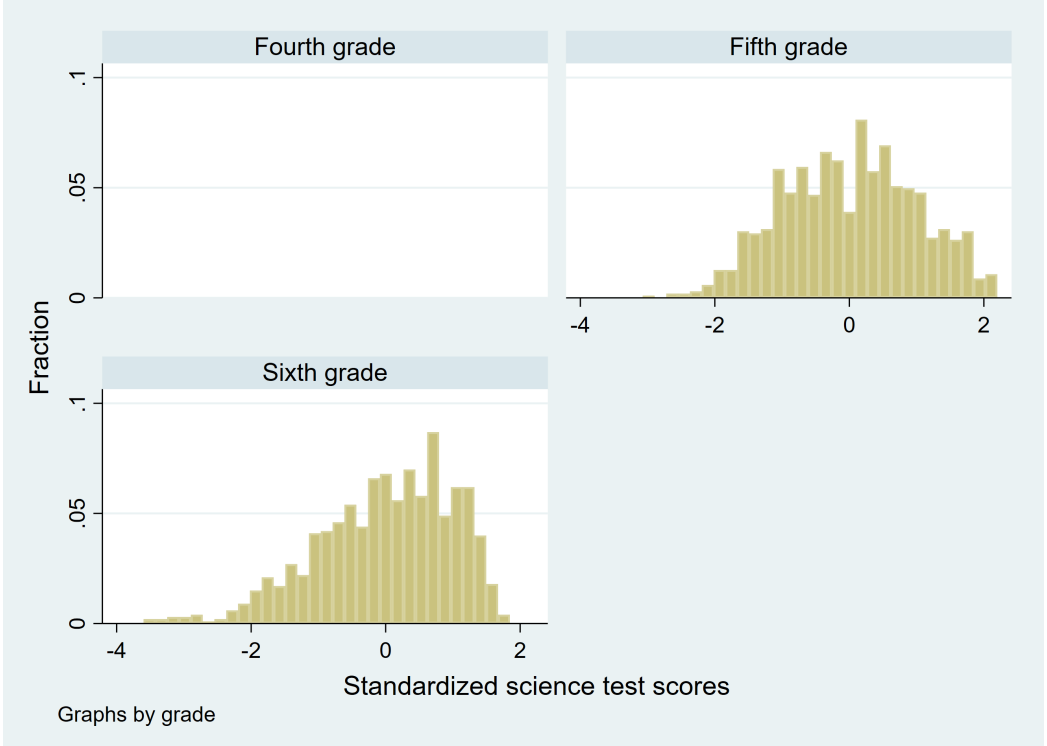

Fourth grade
Fifth grade
Sixth grade
Fraction
0
.05
.1
-4
-2
0
2
Standardized science test scores
Graphs by grade

Figure 2. Distribution of school engagement scores.

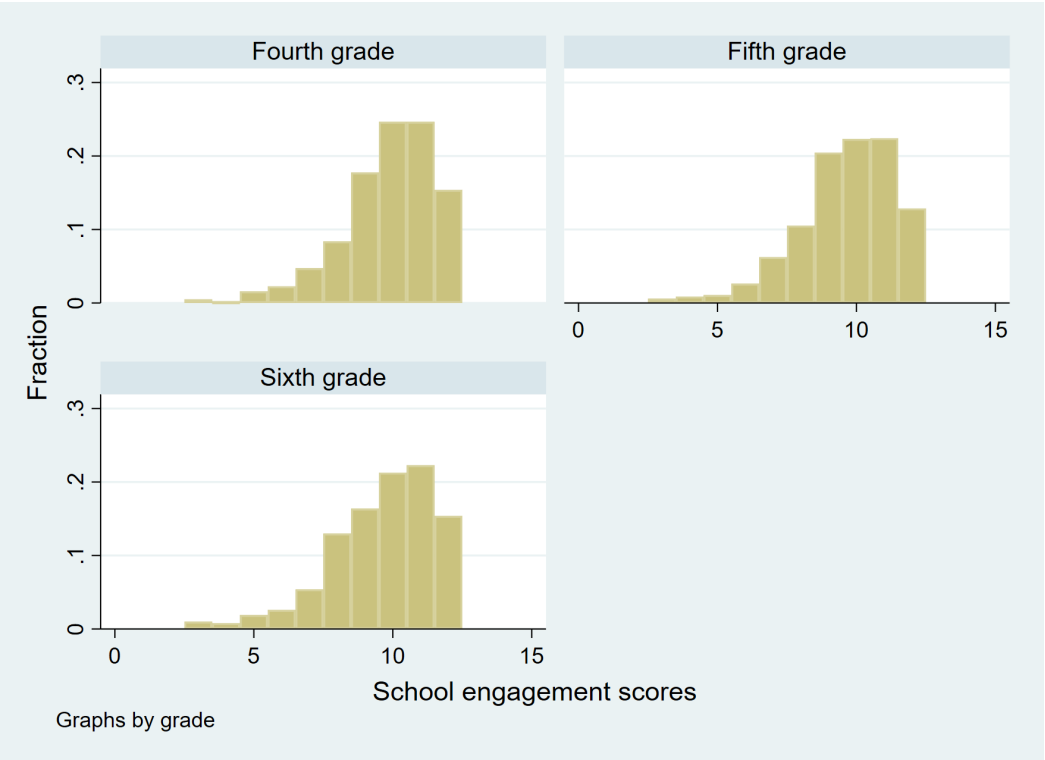


Figure 3. Distribution of friendship outcome scores.

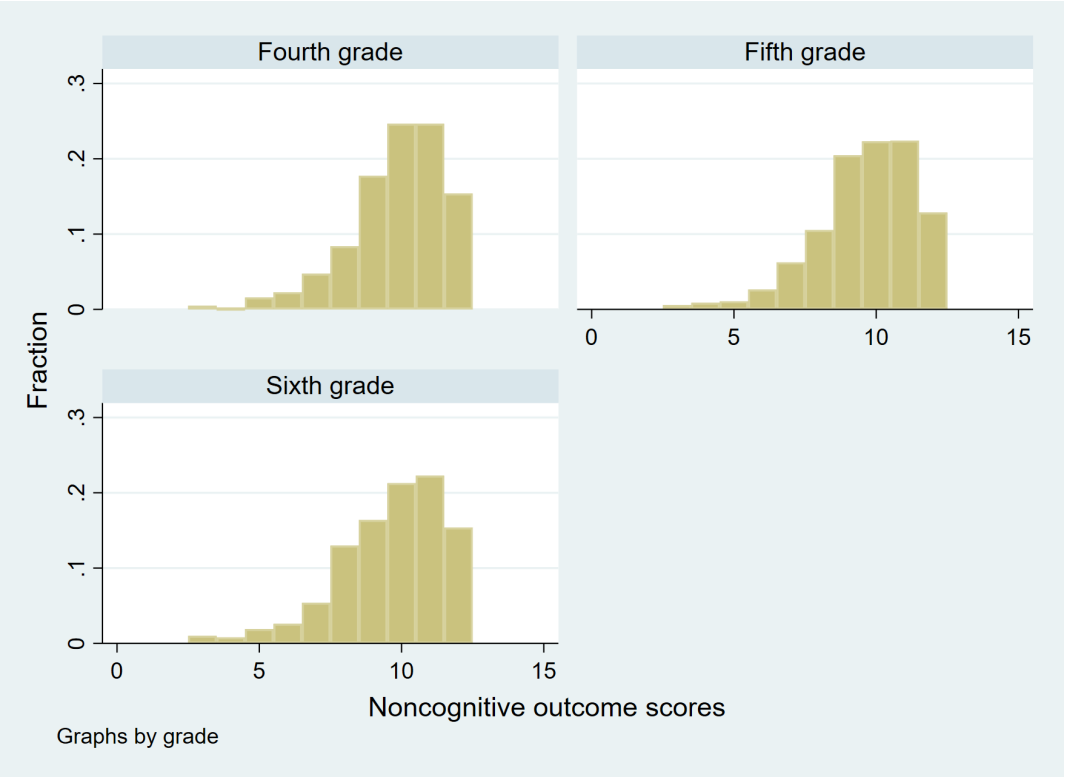


Figure 4. Distribution of annual bullying victimization scores.

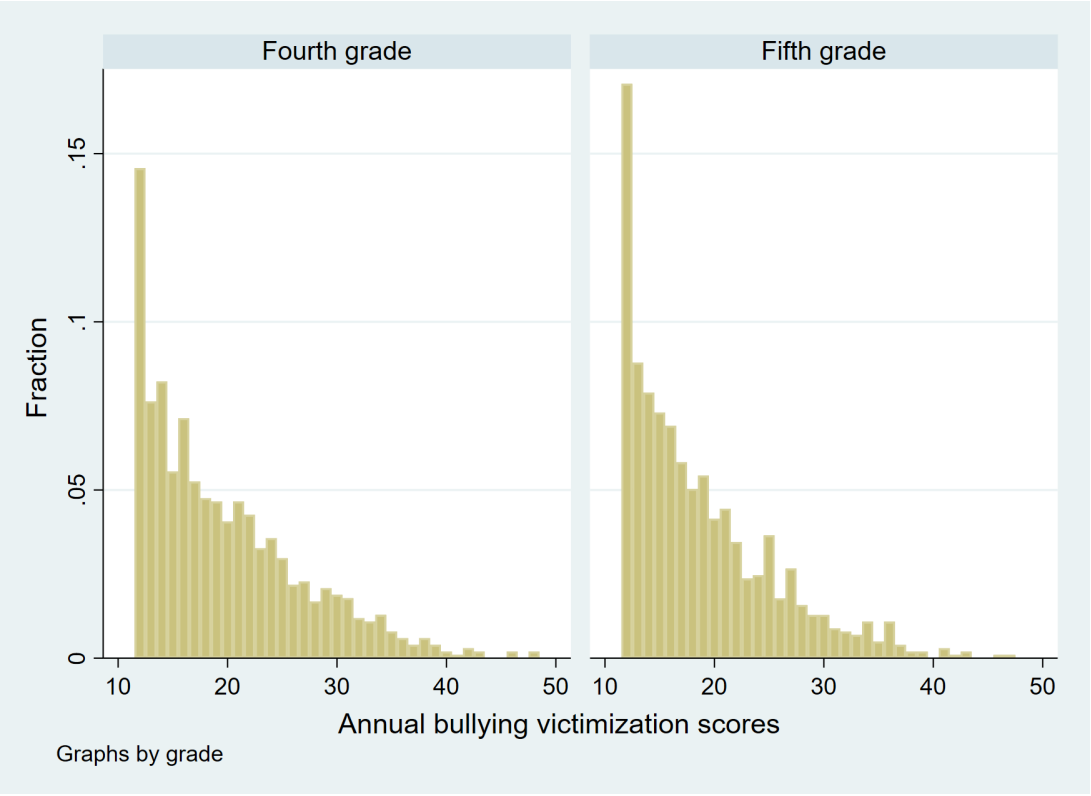


Table 1. Definition of variables.

Variable	Explanation of Variables
Outcome variables:	
Japanese & Math$_{t+1}$	The sum of Japanese and mathematics test scores in period $t+1$, standardized to have a mean of 0 and a standard deviation of 1 for each grade (administered from fourth grade to sixth grade).
Japanese$_{t+1}$	Japanese test scores in period $t+1$, standardized to have a mean of 0 and a standard deviation of 1 for each grade (administered from fourth grade to sixth grade).
Math$_{t+1}$	Mathematics test scores in period $t+1$, standardized to have a mean of 0 and a standard deviation of 1 for each grade (administered from fourth grade to sixth grade).
Sci$_{t+1}$	Science test scores in $t+1$, standardized to have a mean of 0 and a standard deviation of 1 for each grade (administered from fifth grade to sixth grade).
School engagement$_{t+1}$	The sum of the responses to the following three questions from the 1st QU in period $t+1$. Each response is rated on a 4-point Likert scale: 1 “strongly disagree”, 2 “disagree”, 3 “somewhat agree”, and 4 “strongly agree”. The total score is standardized to have a mean of 0 and a standard deviation of 1. 1. Do you feel happy when you can accomplish something at school that you previously could not? 2. Do you enjoy answering the teacher's questions or expressing your opinions during class? 3. Do you strive to achieve good grades and improve your academic abilities?
Friendship$_{t+1}$	The sum of the responses to the following nine questions from the 1st QU in period $t+1$. Each response is rated on a 4-point Likert scale: 1 “strongly disagree”, 2 “disagree”, 3 “somewhat agree”, and 4 “strongly agree”. The total score is standardized to have a mean of 0 and a standard deviation of 1. 1. Do your classmates talk to you and treat you kindly? 2. Are there friends in your class whom you think are kind or admirable? 3. Do you feel that your classmates like you and consider you part of the group? 4. Do your classmates recognize or admire your abilities in areas such as sports, academics, or extracurricular activities? 5. Do your classmates encourage you when you fail? 6. Do you think there is someone in your class who understands your feelings? 7. Do you believe your classmates cooperate with or support you when you try to do something? 8. Do you think there are many people in your class who actively participate in various activities? 9. When you share your thoughts or ideas, do your classmates listen attentively without mocking you?
Regressors of interest:	
Bullied$_{t}$	The sum of responses to the following six questions from both the 1st and 2nd rounds of the QU in period t. Each response is rated on a four-point Likert scale: 1 (Strongly Disagree), 2 (Disagree), 3 (Somewhat Agree), 4 (Strongly Agree). The total score is standardized to have a mean of 0 and a standard deviation of 1. 1. Do you experience distress when someone in your class says unpleasant things to you or makes fun of you? 2. Do you experience distress when someone in your class acts violently toward you? 3. Do you ever feel like you don't want to be in your class because you are mocked or belittled by your classmates? 4. Do you find yourself alone during recess or other breaks? 5. Do you end up being the last one left out when groups are formed in class? 6. Do you ever feel ignored or excluded by your classmates?

$Bullied_{t,1}$	The sum of responses to the six questions above from the first round of the QU in period t.
$Bullied_{t,2}$	The sum of responses to the six questions above from the second round of the QU in period t.
Other explanatory variables:	
Girl	A dummy variable that takes the value 1 if the individual is female, and 0 otherwise.
Class Size	The number of students observed in each class.
Girl Ratio	The proportion of female students in the class.

Table 2. Descriptive statistics.

	Obs.	Mean	SD	Min.	Max.
Outcome variables:					
Japanese & Math$_{t+1}$	2,027	0.00	0.99	-3.45	2.26
Japanese$_{t+1}$	2,027	0.00	0.99	-3.82	2.13
Math$_{t+1}$	2,027	0.00	0.99	-3.24	2.21
Sci$_{t+1}$	1,001	0.00	1.00	-3.60	1.80
School engagement$_{t+1}$	2,027	0.00	1.00	-3.76	1.35
Friendship$_{t+1}$	2,027	0.00	1.00	-4.12	1.44
Regressors of interest:					
Bullied$_{t}$	2,022	0.00	1.00	-1.08	4.28
Bullied$_{t,1}$	2,040	0.00	1.00	-0.98	3.84
Bullied$_{t,2}$	2,022	-0.01	1.00	-0.97	4.13
Other explanatory variables:					
Girl	2,040	0.51	0.50	0.00	1.00
Class Size	2,040	29.2	4.6	13.0	38.0
Girl Ratio	2,040	0.51	0.06	0.30	0.64

Table 3. Determinants of bullying victimization.

	$Bullied_{t,2}$				
	(1)	(2)	(3)	(4)	(5)
$Japanese\ \&\ Math_{t,1}$	-0.1040***			-0.0538**	-0.0479**
	(0.0272)			(0.0250)	(0.0240)
$School\ engagement_{t,1}$		-0.2098***		0.0480*	0.0459*
		(0.0249)		(0.0283)	(0.0273)
$Friendship_{t,1}$			-0.4617***	-0.4806***	-0.4705***
			(0.0274)	(0.0326)	(0.0322)
Girl	-0.0254	0.0113	0.0602	0.0618*	0.0624*
	(0.0418)	(0.0425)	(0.0368)	(0.0369)	(0.0367)
Class size					0.0087
					(0.0057)
Girl Ratio					-0.6709
					(0.4543)
Clas-fixed effects	Yes	Yes	Yes	Yes	No
Obs.	2,022	2,022	2,022	2,022	2,022
R^2	0.092	0.121	0.265	0.269	0.213

Note: Standard errors are clustered at the class level and are reported in parentheses.
Significance levels are denoted as ***1% level, **5% level, *10% level.

Table 4. The effect of bullying on academic performance.

	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$
	(1)	(2)	(3)	(4)
Bullied$_t$	-0.0333**	-0.0520***	-0.0294*	-0.0419*
	(0.0149)	(0.0180)	(0.0165)	(0.0231)
Girl	0.0547**	0.1345***	0.0199	0.0858**
	(0.0242)	(0.0286)	(0.0296)	(0.0404)
Outcome poly. (fifth order, start of year t)	Yes	Yes	Yes	Yes
Class-fixed effects	Yes	Yes	Yes	Yes
Obs.	1,988	1,988	1,988	988
R^2	0.721	0.586	0.630	0.621

Note: Standard errors are clustered at the class level and are reported in parentheses.
Significance levels are denoted as ***1% level, **5% level, *10% level.

Table 5. The effect of bullying on school engagement.

	School Engagement$_{t+1}$
Bullied$_t$	-0.1066***
	(0.0276)
Girl	0.0298
	(0.0390)
Outcome poly. (fifth order, start of year t)	Yes
Class-fixed effects	Yes
Obs.	1,992
R^2	0.326

Note: Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table 6. The effect of bullying on friendship outcomes.

	$Friendship_{t+1}$
$Bullied_t$	-0.2290***
	(0.0274)
Girl	0.0889**
	(0.0359)
Outcome poly. (fifth order, start of year t)	Yes
Class-fixed effects	Yes
Obs.	1,992
R^2	0.407

Note: Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table 7. Heterogenous effects.

	Japanese & Math$_{t+1}$		Japanese$_{t+1}$		Math$_{t+1}$		Sci$_{t+1}$		School engagement$_{t+1}$		Friendship$_{t+1}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Bullied$_t$	-0.0466**	-0.0276	-0.0384*	-0.0369*	-0.0637***	-0.0348*	-0.0981***	-0.0365	-0.0677*	-0.1173***	-0.2054***	-0.2631***
	(0.0180)	(0.0178)	(0.0222)	(0.0196)	(0.0207)	(0.0199)	(0.0305)	(0.0309)	(0.0342)	(0.0351)	(0.0353)	(0.0310)
×Class size_large	0.0252		-0.0257		0.0653**		0.1024**		-0.0727		-0.0462	
	(0.0290)		(0.0345)		(0.0324)		(0.0424)		(0.0516)		(0.0400)	
×Girl ratio_high		-0.0137		-0.0360		0.0130		-0.0127		0.0255		0.0842**
		(0.0300)		(0.0360)		(0.0340)		(0.0467)		(0.0525)		(0.0395)
Girl	0.0555**	0.0547**	0.1337***	0.1345***	0.0219	0.0197	0.0864**	0.0855**	0.0272	0.0299	0.0879**	0.0884**
	(0.0242)	(0.0242)	(0.0285)	(0.0285)	(0.0295)	(0.0296)	(0.0402)	(0.0402)	(0.0390)	(0.0390)	(0.0357)	(0.0355)
Outcomes poly. (fifth order, start of year)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,988	1,988	1,988	1,988	1,988	1,988	988	988	1,992	1,992	1,992	1,992
R^2	0.722	0.721	0.586	0.586	0.631	0.630	0.623	0.621	0.327	0.326	0.408	0.409

Note: Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table 8. The effects of peer-bullying victimization on student outcomes.

	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$	School Engagement$_{t+1}$	Friendship$_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
Peer Bullied$_t$	-0.1467*	-0.1329	-0.1193	-0.2253	-0.1524	-0.0836
	(0.0853)	(0.0930)	(0.1208)	(0.1713)	(0.0991)	(0.0967)
Bullied$_t$	-0.0377**	-0.0549***	-0.0352**	-0.0445*	-0.1073***	-0.2334***
	(0.0149)	(0.0176)	(0.0168)	(0.0225)	(0.0268)	(0.0268)
Girl	0.0386	0.1178***	0.0086	0.0694*	0.0106	0.0753*
	(0.0258)	(0.0303)	(0.0302)	(0.0396)	(0.0398)	(0.0404)
Class-level peer attributes	Yes	Yes	Yes	Yes	Yes	Yes
Outcomes poly. (fifth order, start of year)	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,988	1,988	1,988	988	1,992	1,992
R^2	0.692	0.555	0.592	0.586	0.280	0.364

Note: Class-level peer attributes include peers' prior cognitive, school engagement and friendship outcomes, the proportion of female students, and class size. Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Appendices

Table A1. Implementation schedule of academic assessments and QU.

	April	May or June	November or December
2014 (fourth grade)	Academic Assessment	QU 1st Round	QU 2nd Round
2015 (fifth grade)	Academic Assessment	QU 1st Round	QU 2nd Round
2016 (sixth grade)	Academic Assessment	QU 1st Round	

Table A2. Data structure.

Academic year	Grade	Student count from city statistics	Our data		
			Student count after dropping those with missing values	Number of schools	Number of classes
2014	4	1,062	1,014	16	38
2015	5	1,054	1,026	16	37
2016	6	1,049	1,001	16	35

Table A3. Additional specifications.

	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$	School engagement$_{t+1}$		Friendship$_{t+1}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ΔBullid$_t$	-0.0164	-0.0356*	-0.0059	-0.0143	-0.0783***		-0.1939***	
	(0.0169)	(0.0191)	(0.0206)	(0.0252)	(0.0272)		(0.0290)	
ΔBullid$_{t\rightarrow t+1}$						-0.0524***		-0.1270***
						(0.0074)		(0.0089)
Bullied$_{t,1}$	-0.0384**	-0.0605***	-0.0334*	-0.0470*	-0.1242***	-0.2040***	-0.2645***	-0.4613***
	(0.0172)	(0.0205)	(0.0193)	(0.0261)	(0.0323)	(0.0323)	(0.0320)	(0.0303)
Girl	0.0547**	0.1340***	0.0203	0.0850**	0.0286	0.0320	0.0841**	0.0897***
	(0.0242)	(0.0286)	(0.0295)	(0.0402)	(0.0391)	(0.0383)	(0.0365)	(0.0327)
Outcome poly. (fifth order, start of	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,988	1,988	1,988	988	1,992	2,002	1,992	2,002
R^2	0.721	0.586	0.630	0.621	0.326	0.345	0.413	0.528

Note: ΔBullying$_t$ is the within-year change in bullying victimization, computed as Bullied$_{t,2}$ minus Bullied$_{t,1}$ in school year t. ΔBullying$_{t\rightarrow t+1}$ is the cross-year change, computed as Bullied$_{t+1,1}$ minus Bullied$_{t,1}$. All difference measures are standardized in the estimation sample. Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table A4. Robustness checks: Effects of bullying on educational outcomes with additional controls.

	Japanese & Math$_{t+1}$			School Engagement$_{t+1}$			Friendship$_{t+1}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Bullied$_t$	-0.0333**	-0.0317**	-0.0316**	-0.1066***	-0.1054***	-0.1054***	-0.2290***	-0.2299***	-0.2302***
	(0.0149)	(0.0149)	(0.0149)	(0.0276)	(0.0268)	(0.0268)	(0.0274)	(0.0260)	(0.0261)
Girl	0.0547**	0.0538**	0.0542**	0.0298	0.0280	0.0279	0.0889**	0.0878**	0.0877**
	(0.0242)	(0.0244)	(0.0244)	(0.0390)	(0.0400)	(0.0401)	(0.0359)	(0.0359)	(0.0359)
Outcomes poly. (fifth order, start of year)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-fixed effects (t)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class-fixed effects (t+1)	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Teacher-fixed effects	No	No	Yes	No	No	Yes	No	No	Yes
Obs.	1,988	1,988	1,988	1,992	1,992	1,992	1,992	1,992	1,992
R^2	0.721	0.729	0.729	0.326	0.348	0.348	0.407	0.431	0.431

Note: Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table A5. Robustness checks: Effects of early- and late-year bullying victimization on educational outcomes.

	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$	School Engagement$_{t+1}$	Friendship$_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
Bullied$_{t,1}$	-0.0302**	-0.0421**	-0.0309**	-0.0423*	-0.0900***	-0.1519***
	(0.0146)	(0.0180)	(0.0150)	(0.0246)	(0.0260)	(0.0309)
Girl	0.0556**	0.1381***	0.0182	0.0817*	0.0313	0.0794**
	(0.0245)	(0.0290)	(0.0295)	(0.0406)	(0.0386)	(0.0359)
Outcomes poly. (fifth order, start of year)	Yes	Yes	Yes	Yes	Yes	Yes
Class-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,997	1,997	1,997	993	2,002	2,002
R^2	0.720	0.584	0.630	0.618	0.321	0.386
	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$	School Engagement$_{t+1}$	Friendship$_{t+1}$
	(7)	(8)	(9)	(10)	(11)	(12)
Bullied$_{t,2}$	-0.0292*	-0.0503***	-0.0217	-0.0348	-0.1048***	-0.2291***
	(0.0150)	(0.0173)	(0.0179)	(0.0214)	(0.0273)	(0.0272)
Girl	0.0539**	0.1333***	0.0195	0.0865**	0.0266	0.0803**
	(0.0243)	(0.0286)	(0.0297)	(0.0400)	(0.0392)	(0.0364)
Outcomes poly. (fifth order, start of year)	Yes	Yes	Yes	Yes	Yes	Yes
Class-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,988	1,988	1,988	988	1,992	1,992
R^2	0.721	0.585	0.630	0.620	0.326	0.412

Note: Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table A6. Estimates on the subsample with item-level QU responses.

	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$	School Engagement$_{t+1}$	Friendship$_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(7)
Bullid$_t$	-0.0336	-0.0402*	-0.0447*	-0.0216	-0.0793**	-0.2175***
	(0.0205)	(0.0215)	(0.0249)	(0.0331)	(0.0366)	(0.0397)
Girl	0.0537	0.1469***	-0.0047	0.0120	0.0282	0.1451***
	(0.0346)	(0.0367)	(0.0423)	(0.0466)	(0.0603)	(0.0509)
Outcome poly. (fifth order, start of year t)	Yes	Yes	No	Yes	No	No
Obs.	1,038	1,038	1,038	580	1,041	1,041
R^2	0.723	0.573	0.649	0.616	0.348	0.410

Note: Analysis restricted to the subsample with item-level QU responses. Standard errors are clustered at the school level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table A7. Exploratory factor analysis of bullying items.

Item	Title	One factor	Two factors	
			Factor 1	Factor 2
1	Distress when teased/mocked	**0.879**	**0.894**	0.032
2	Distress when violence	**0.786**	**0.853**	-0.047
3	Don't want to be in class (mocked/belittled)	**0.881**	**0.751**	0.176
4	Alone at recess/breaks	**0.623**	0.008	**0.802**
5	Last one left out in groups	**0.625**	0.020	**0.786**
6	Ignored/excluded by classmates	**0.809**	**0.429**	**0.499**
eigenvalue		3.599	3.693	0.574

Note: This table reports factor loadings from exploratory factor analysis (EFA). The EFA is based on the decomposition of the polychoric correlation matrix and uses oblique Promax. Analysis restricted to the subsample with item-level QU responses.

Tabel A8. Confirmatory actor analysis fit statistics for one- and two-factor bullying-victimization models.

Model	RMSEA	CFI	TLI
One factor	0.179	0.902	0.837
Two factors	0.116	0.964	0.932

Note: The table reports fit indices for CFA models. Analysis restricted to the subsample with item-level QU responses. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker–Lewis Index.

Tabel A9. Robustness checks: Effects of bullying victimization (CFA-1 factor score) on educational outcomes.

	Japanese & $Math_{t+1}$	$Japanese_{t+1}$	$Math_{t+1}$	Sci_{t+1}	School $Engagement_{t+1}$	$Friendship_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
$Bullying\ (CFA\text{-}1)_t$	-0.0341*	-0.0403**	-0.0457*	-0.0204	-0.0728*	-0.2125***
	(0.0198)	(0.0195)	(0.0249)	(0.0322)	(0.0374)	(0.0416)
Girl	0.0533	0.1465***	-0.0053	0.0120	0.0267	0.1408***
	(0.0345)	(0.0366)	(0.0423)	(0.0465)	(0.0604)	(0.0513)
Outcome poly. (fifth order, start of year t)	Yes	Yes	No	Yes	No	No
Obs.	1,038	1,038	1,038	580	1,041	1,041
R^2	0.723	0.573	0.649	0.616	0.347	0.410

Note: Analysis restricted to the subsample with item-level QU responses.$Bullying\ (CFA\text{-}1)_t$ denotes the standardized bullying-victimization factor score from the one-factor confirmatory factor model discussed in Section 6. Standard errors are clustered at the school level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Tabel A10. Robustness checks: Effects of bullying victimization (CFA-2 factor scores) on educational outcomes.

	Japanese & Math$_{t+1}$		Japanese$_{t+1}$		Math$_{t+1}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Bullid (CFA-2, direct)$_t$	-0.0326		-0.0380*		-0.0446*	
	(0.0195)		(0.0190)		(0.0248)	
Bullid (CFA-2, indirect)$_t$		-0.0356*		-0.0423*		-0.0457*
		(0.0198)		(0.0212)		(0.0243)
Girl	0.0530	0.0543	0.1462***	0.1477***	-0.0057	-0.0039
	(0.0345)	(0.0347)	(0.0366)	(0.0368)	(0.0423)	(0.0425)
Outcome poly. (fifth order, start of year t)	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1,038	1,038	1,038	1,038	1,038	1,038
R^2	0.723	0.723	0.573	0.574	0.648	0.649
	Sci$_{t+1}$		School Engagement$_{t+1}$		Friendship$_{t+1}$	
	(7)	(8)	(9)	(10)	(11)	(12)
Bullid (CFA-2, direct)$_t$	-0.0161		-0.0706*		-0.2026***	
	(0.0325)		(0.0378)		(0.0431)	
Bullid (CFA-2, indirect)$_t$		-0.0310		-0.0744**		-0.2110***
		(0.0306)		(0.0351)		(0.0358)
Girl	0.0122	0.0122	0.0255	0.0295	0.1362**	0.1491***
	(0.0464)	(0.0467)	(0.0605)	(0.0602)	(0.0518)	(0.0504)
Outcome poly. (fifth order, start of year t)	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	580	580	1,041	1,041	1,041	1,041
R^2	0.616	0.617	0.347	0.347	0.408	0.409

Note: Analysis restricted to the subsample with item-level QU responses. Bullying (CFA-2, direct) t and Bullying (CFA-2, indirect) t denote standardized factor scores from the two-factor confirmatory factor model representing direct and indirect bullying, respectively, as discussed in Section 6. Standard errors are clustered at the school level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table A11. Exploratory factor analysis of school engagement.

Item	Title	One factor
1	Positive affect after mastery	**0.781**
2	Enjoy answering in class	**0.614**
3	Strive for good grades	**0.700**
eigenvalue		1.476

Note: This table reports factor loadings from exploratory factor analysis (EFA). The EFA is based on the decomposition of the polychoric correlation matrix and uses oblique Promax. Analysis restricted to the subsample with item-level QU responses.

Table A12. Exploratory factor analysis of friendship outcomes.

Item	Title	One factor
1	Treated kindly by classmates	**0.810**
2	Kind or admirable friends	**0.695**
3	Liked and part of the group	**0.763**
4	Abilities recognized or admired	**0.644**
5	Encouraged when you fail	**0.865**
6	Someone understands your feelings	**0.787**
7	Classmates cooperate and support	**0.860**
8	Many classmates actively participate	**0.608**
9	Listen attentively without mocking	**0.701**
eigenvalue		5.105

Note: This table reports factor loadings from exploratory factor analysis (EFA). The EFA is based on the decomposition of the polychoric correlation matrix and uses oblique Promax. Analysis restricted to the subsample with item-level QU responses.

Table A13. Robustness checks: Effects of bullying on school engagement (CFA-1 factor score) and friendship Outcomes (CFA-1 factor score).

	School Engagement $(CFA\text{-}1)_{t+1}$	Friendship $(CFA\text{-}1)_{t+1}$
	(1)	(2)
$Bullid_t$	-0.0864**	-0.2112***
	(0.0374)	(0.0358)
Girl	0.0627	0.1417***
	(0.0581)	(0.0504)
Outcome poly. (fifth order, start of year t)	Yes	Yes
Obs.	1,021	1,021
R^2	0.343	0.417

Note: Analysis restricted to the subsample with item-level QU responses.School Engagement $(CFA\text{-}1)_{t+1}$ and Friendship $(CFA\text{-}1)_{t+1}$ denote standardized factor scores from one-factor confirmatory factor models of the respective item sets. Standard errors are clustered at the school level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table A14. Estimates using a victimization index excluding social-isolation items.

	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$	School Engagement$_{t+1}$	Friendship$_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
Bullid$_t$ (excl.	-0.0292	-0.0368**	-0.0403	-0.0259	-0.0675*	-0.1966***
isolation-related	(0.0190)	(0.0179)	(0.0247)	(0.0323)	(0.0369)	(0.0432)
Girl	0.0520	0.1310***	0.0013	-0.0069	0.0229	0.1197**
	(0.0339)	(0.0385)	(0.0425)	(0.0496)	(0.0596)	(0.0513)
Outcome poly. (fifth order, start of year t)	Yes	Yes	No	Yes	No	No
Obs.	1,057	1,057	1,057	599	1,060	1,060
R^2	0.723	0.571	0.645	0.608	0.351	0.405

Note: Analysis restricted to the subsample with item-level QU responses.The bullying victimization index excludes items that primarily capture social isolation. Standard errors are clustered at the school level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.

Table A15. Robustness checks: Mean reversion and unobserved selection.

	Japanese & Math$_{t+1}$		Japanese$_{t+1}$		Math$_{t+1}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Bullied$_t$	-0.1390***	-0.0333**	-0.1315***	-0.0520***	-0.1262***	-0.0294*
	(0.0230)	(0.0149)	(0.0235)	(0.0180)	(0.0227)	(0.0165)
Girl	0.1956***	0.0547**	0.3793***	0.1345***	0.0070	0.0199
	(0.0395)	(0.0242)	(0.0418)	(0.0286)	(0.0399)	(0.0296)
Outcomes poly. (fifth order, start of year)	No	Yes	No	Yes	No	Yes
Class-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1988	1988	1988	1988	1988	1988
R^2	0.092	0.721	0.106	0.586	0.087	0.630
Oster's δ ($R^2_{max} = 1.3R^2$)		0.913		1.761		0.869
	Sci$_{t+1}$		School Engagement$_{t+1}$		Friendship$_{t+1}$	
	(7)	(8)	(9)	(10)	(11)	(12)
Bullied$_t$	-0.0882***	-0.0415*	-0.2258***	-0.1066***	-0.4548***	-0.2290***
	(0.0263)	(0.0231)	(0.0264)	(0.0276)	(0.0225)	(0.0274)
Girl	0.0411	0.0851**	0.1507***	0.0298	0.1791***	0.0889**
	(0.0666)	(0.0401)	(0.0465)	(0.0390)	(0.0412)	(0.0359)
Outcomes poly. (fifth order, start of year)	No	Yes	No	Yes	No	Yes
Class-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	988	988	1992	1992	1992	1992
R^2	0.087	0.621	0.137	0.326	0.300	0.407
Oster's δ ($R^2_{max} = 1.3R^2$)		2.532		1.626		0.900

Note: Standard errors are clustered at the class level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level. Oster's δ follows Oster (2019) and quantifies how strong selection on unobservables would need to be, relative to observables, to reduce the estimated bullying effect to zero. Following Oster (2019), we set $R^2_{\max}$ to 1.3 times the R^2 from the full specification. To compute Oster's δ, we use the Stata ado-file psacalc provided by Emily Oster.

Table A16. Class-level bullying victimization and outcomes.

	Japanese & Math$_{t+1}$	Japanese$_{t+1}$	Math$_{t+1}$	Sci$_{t+1}$	School Engagement$_{t+1}$	Friendship$_{t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
Class-level bullying$_t$	-0.1988**	-0.1809*	-0.1970**	-0.1112	-0.0894	-0.2626**
	(0.0702)	(0.0854)	(0.0836)	(0.1560)	(0.0805)	(0.1189)
Girl	0.0547**	0.1340***	0.0203	0.0850**	0.0286	0.0841**
	(0.0242)	(0.0286)	(0.0295)	(0.0402)	(0.0391)	(0.0365)
Outcome poly. (fifth order, start of year t)	Yes	Yes	No	Yes	No	No
Obs.	75	75	75	37	75	75
R^2	0.529	0.470	0.450	0.578	0.455	0.614

Note: Estimates from regressions of class-average outcomes in year $t+1$ on class-level bullying victimization in year t. Standard errors are clustered at the school level and are reported in parentheses. Significance levels are denoted as ***1% level, **5% level, *10% level.