CENTER FOR GLOBAL DEVELOPMENT

Improving School Management of Violence: Evidence from a Nationwide Policy in Peru

Gabriela Smarrelli

Abstract

Exposure to school violence has been proven to be detrimental to human capital formation, but there is limited rigorous evidence about how to tackle this pervasive issue. This paper examines the impacts of a large-scale government intervention that aimed to improve school leaders' skills to manage school violence in Peru. I exploit the eligibility rules used to select beneficiary schools and use a fuzzy regression discontinuity design to estimate the short-term impacts of the intervention on violence and education-related outcomes. The findings show that the likelihood of reporting violence increased by 15 percentage points and that the number of reports of violence rose among eligible schools. Combining unique administrative and primary data, I provide suggestive evidence that the documented rise in reports of violence is primarily due to shifts in reporting rather than a greater incidence of school violence. Upon exploring the short-term impacts on education-related outcomes, I find the intervention reduced students' likelihood of switching schools by two percentage points. These findings add to our understanding of the benefits of investing in school staff skills for safer learning environments.

KEYWORDS
economics of education, school management of violence, school mobility, school dropout, test scores

JEL CODES
I20, I29, H75

WORKING PAPER 667 · NOVEMBER 2023

Improving School Management of Violence: Evidence from a Nationwide Policy in Peru

Gabriela Smarrelli
Center for Global Development
Email: gsmarrelli@cgdev.org

I am grateful to Alison Andrew, Diala Al Masri, Antonella Bancalari, Lee Crawfurd, David Evans, Douglas Gollin, Susannah Hares, Pramila Krishnan, Sonya Krutikova, Clare Leaver, Justin Sandefur, Abhijeet Singh, Marcos Vera-Hernandez, Christopher Woodruff, and the participants at workshops within Oxford University and at the RISE, CSAE, CESifo, RES and EEA Conferences for their comments, guidance and stimulating discussions. Many thanks as well to the Peruvian Ministry of Education, particularly to those at the Dirección General de Calidad de la Gestión Escolar, Unidad de Seguimiento y Evaluación, and Unidad de Estadística. Special thanks to the UGEL School Coexistence Specialists and the school principals that participated in the qualitative interviews and the online surveys. Every conversation, meeting and interaction were essential to developing this research. I am also grateful to the team of survey administrators. Without the support of Ariana Cisneros, Ana Lucia Andrade, Gabriela Cueva, Sofia Plasencia, and Pedro Ramos, data collection would not have been possible. I would also like to thank the Economic and Social Research Council and Pousaz Philanthropies for the financial support. My gratitude extends to the unmentioned individuals that supported me throughout this research project.

Gabriela Smarrelli. 2023. “Improving School Management of Violence: Evidence from a Nationwide Policy in Peru.” CGD Working Paper 667. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/improving-school-management-violence-evidence-nationwide-policy-peru

CENTER FOR GLOBAL DEVELOPMENT

2055 L Street, NW Fifth Floor

Washington, DC 20036

202.416.4000

1 Abbey Gardens

Great College Street

London

SW1P 3SE

www.cgdev.org

1 Introduction

Almost a third of students aged 13 to 15 years worldwide have been victims of violence in schools (UNESCO 2019). Violence can take multiple forms - including physical, sexual, and psychological violence - and can emerge as a set of isolated events or as repeated attacks against the same victim. The prevalence of intimidation and abuse in schools and the lack of actions to address it directly affects children's right to inclusive and equitable education and is detrimental to human capital formation.

Extensive evidence, mainly from the educational psychology literature and to a lesser extent from the economics literature, has found a negative association between being a victim of school violence and learning outcomes (Ponzo 2013; Strøm et al. 2013; Eriksen et al. 2014; Contreras et al. 2016; Delprato et al. 2017), as well as a positive association with student dropout, student mobility, and absenteeism (Brown and Taylor 2008; Dunne et al. 2013; Carson et al. 2013; Burdick-Will et al. 2021). Evidence also documents lasting negative effects over the life cycle both in the likelihood of employment during adulthood (Varhama and Björkqvist 2005; Brown and Taylor 2008), and dimensions of individual wellbeing related to mental health (Kim et al. 2005; Hinduja and Patchin 2010; Hepburn and Miller 2012; Sarzosa and Urzúa 2021). Adverse effects that extend to the perpetrators of violence (Wolke et al. 2013; Wolke and Lereya 2015) and the bystanders (Rivers et al. 2009).

The negative consequences of school-based violence and the recognition of this phenomenon as a public health issue have led to a rise in laws, policies and programs targeting the school safety of children (Kelly 2017; Rees et al. 2022; Chávez et al. 2020). Since the early 2000's the United States has implemented state-specific anti-bullying laws and the United Kingdom has enacted the Education and Inspection Act to address the issue of school violence. In the last decade, 16 of 33 countries in Latin America and the Caribbean have also enacted laws to protect children against school violence. Moreover, governments and non-government organizations have implemented school-specific violence prevention programs.[1] However, partly due to empirical challenges related to the absence of valid comparison groups and the scarce availability of data from either violence reports or victimization surveys, there is limited rigorous causal evidence on how to effectively address violence in schools, particularly in low and middle-income settings and for large-scale interventions.

In this paper, I study whether investing in school principals' skills to manage school violence represents an avenue to improve student experiences in school. I focus on the context of Peru, a middle-income country with one of the highest rates of victimization at schools in Latin America, and analyze a large-scale technical assistance (TA) program designed in 2019 by the Ministry of Education (MINEDU) to improve the management of violence in public schools. I exploit the eligibility rules used by MINEDU to select the beneficiary schools and, using a unique administrative dataset at the school and student level, I study the short-term impacts of the TA on violence-related and education-related outcomes.

The TA consisted of 3 cycles of training activities directed to school principals. The training topics included the identification and monitoring of different forms of violence (physical, psychological

[1] See Appendix A for details.

and sexual) between students and teacher-to-student, the adoption of response protocols, and the implementation of positive discipline strategies. In each of the provinces of the country, MINEDU offered the intervention to 12 schools: 3 nucleo schools and 9 adjacent schools.[2] In this paper, I limit the analysis to the adjacent schools mainly for two reasons. First, 90 percent of nucleo schools were targeted to receive another intervention at the end of the school year, making it harder to disentangle the effect of the TA. Second, the eligibility rules for adjacent schools allow me to study the impact of the program using a fuzzy regression discontinuity design (RDD) that is likely to produce more credible estimates.

To select the adjacent schools, MINEDU chose, in each province of the country, the top-9 schools that were closest in distance to one of the 3 nucleo schools and that had the highest number of enrolled students. I exploit the fact that the eligibility criterion mimics a ranking procedure where the schools in the top-9 of the ranking were assigned to receive the intervention and those just above the threshold rule of 9 were not because they were a few kilometres further away from the nucleo schools and/or because they had a lower number of enrolled students.

MINEDU only kept a record of the schools that were targeted to receive the TA. Therefore, based on the selection criteria, I re-create the ranking of schools and generate an eligibility dummy that takes the value of 1 for those schools located below the threshold of 9; and 0 otherwise. I show graphically that there are discontinuities in the probability of treatment. I produce intention to treat estimates and, using the eligibility dummy as an instrument for treatment, I produce instrumental variable estimates that provide the local average treatment effect on violence-related indicators including the likelihood of reporting violence and the number of reports of violence, and education-related indicators including student dropout, student mobility, and test scores.

A Peruvian platform called SíSeVE allowed me to observe all the reported incidents of school violence. Using this novel dataset, I find that the intervention increased the likelihood of reporting by 15 percentage points. I also observe that the number of reports of violence increased, on average, by 1 report. This increase is non-trivial considering that, conditional on reporting, the mean and median number of reports of violence among the comparison schools was 3 and 2 reports, respectively.

One plausible explanation is that these results reflect a change in reporting behaviour. Being a victim, a confidant of the victim or a witness of school violence will not necessarily translate into reporting an incident. The absence or lack of knowledge about the available channels for reporting and the uncertainty about the school's ability to deal with violence, coupled with feelings of shame and guilt, fear of retaliation from the perpetrator, and fear of disapproval from social networks constrain the decision of reporting violence (Skogan 1984; Cortes and Kochenderfer-Ladd 2014; Xie and Baumer 2019). Administrative and qualitative data suggest that the intervention could have contributed to reducing barriers to reporting. I observe that eligible schools worked on more practices related to the management of school violence. Examples of practices include the creation of spaces to discuss the

[2]This categorization of schools was specific to the intervention. Some beneficiary schools were designated 'nucleo' schools (a Spanish word for centre or nucleus) because the group learning sessions of the training were carried out in them. Leaders from nearby schools, designated 'adjacent,' travelled to nucleo schools to attend these sessions.

topic of school violence, the dissemination of information about where to report incidents of violence, the development of school coexistence rules and the incorporation of the topic of school violence in classroom plans. These practices might have contributed to reducing uncertainty and fears related to reporting by reducing information barriers about where to report and by increasing trust in the school's ability to address incidents of school violence.

An alternative story could be that the increase in reporting reflects an increase in violence levels. This story is perhaps more difficult to study considering the issue of underreporting. It is widely accepted in the literature that data based on violence reports does not necessarily reflect the true prevalence of violence. Empirical research studying different forms of violence, including domestic and sexual violence, has shown that compared to survey-victimization data, report-based data generally underestimates victimization rates (Skogan 1984; García-Moreno 2005; Doleac and Carr 2016; Xie and Baumer 2019). In the Peruvian context, the evidence suggests that not every event of violence is being reported and points to the importance of reducing the barriers to reporting. For instance, administrative data shows that in 90 percent of the schools that documented a case of violence for the first time in 2019, students had already witnessed events of physical and psychological violence before but did not report them.

Using this data, I create an index of perceptions of school violence where higher values indicate that the student witnessed a higher number of events of physical and psychological violence in the school. The data shows no significant difference in the index of perception of school violence between the schools eligible to receive the intervention and those that were not, providing suggestive evidence that violence levels did not increase due to the intervention. Instead, in the short term, my estimates are more likely to reflect shifts in reporting behaviour. This finding is particularly relevant for policymakers as underreporting limits the possibility of dealing with and reducing future events of violence.[3]

Next, I turn to education-related outcomes. Using student-level data, I study the impact of the program on student dropout, student mobility (i.e., switching schools), and test scores. I find that the intervention did not have an impact on student dropout, but it reduced the likelihood of student mobility by around 2 percentage points, which corresponds to a 20 percent reduction in student mobility relative to the comparison group. The indicator of student mobility only considers non-structural moves that occur when the student could, in theory, have stayed at their previous school. A common reason for these types of moves is residential mobility. That is, cases in which the family might move to a new province (e.g., due to divorce or carer's access to a new job) and consequently, the student switches schools (Welsh, 2017). I consider this in the analysis and observe that the estimated coefficient is not driven by changes in residential mobility, but instead is likely to be explained by changes in the student's experience of school.

Moreover, the intervention did not have an impact on learning outcomes. Test scores are measured

[3]This evidence relates to the work by Iyer et al. (2012) that study the relationship between female representation in local governments and crime in India. Even though it is a different field of study, the authors also show that the rise in documented crimes in villages with higher female representation is driven by greater reporting rather than actual increases in crime.

one month after the intervention was completed so it may be too soon to measure impacts on these variables. Unfortunately, the lack of data on test scores after 2020 means I cannot explore whether the intervention had medium-term effects on learning.

Taken together, my findings suggest that the intervention changed the reporting outlet and reduced the likelihood of student mobility across schools. Drawing on the voice and exit framework (Hirschman 1972), these findings can be interpreted to reflect the benefits of providing a space to communicate or speak up about experiences of violence. Being able to report cases of violence and observing that the school takes actions to address violence may have improved students' experiences of school and their psychological well-being, influencing the student's decision to stay in the same school, even if actual levels of violence may have not changed in the short term.

The results are robust to changes in the functional form, the estimates remain similar at different windows of analysis, and I do not observe jumps in any of the pre-treatment outcome variables. Therefore, it is unlikely there are serious threats to the internal validity of the estimates. However, the empirical strategy has two main limitations that are frequent in RDD settings as the method only uses a sample around the threshold. First, my estimates are less likely to be relevant for the schools located far from the threshold. These are schools that are more likely to be rural and with a smaller number of enrolled students. Second, the statistical power to detect heterogeneity is limited, restricting the possibility of analysing with precision whether the impacts differ by pre-treatment characteristics.

School violence is a widespread and persistent problem. The return to in-person instruction after the pandemic, showed increases in bullying to pre-pandemic levels, signalling the persistent nature of bullying and violence dynamics when no actions are taken to tackle it (Bacher-Hicks et al. 2022). This paper contributes to the scarce literature studying how to address violence in schools (Kelly 2017; Chávez et al. 2020). The limited rigorous available evidence comes mainly from high-income countries that have studied the impact of state antibullying laws (Rees et al. 2022), or the effects of school-specific interventions. These interventions follow either a 'student-only approach' that focuses on students' skill development as a mechanism to prevent school violence or a 'whole-school approach' that incorporates school staff training components.[4] Both types of interventions have been found to reduce the likelihood of student victimization in high-income settings (Olweus 2005; Kärnä et al. 2011; Limber et al. 2018; Nocentini and Menesini 2016; Espelage et al. 2013; Bradshaw et al. 2015). However, we still lack knowledge regarding their efficacy in contexts that differ substantially from high-income countries.

Low- and middle-income countries face different constraints related to the quality of the systems of education, the budget, the social norms on violence and the culture of punitive discipline that prevails in many settings. To my knowledge, other than this paper, only a few interventions have been rigorously studied. In Uganda, Devries et al. (2015) and Knight et al. (2018) found that a whole-school intervention called 'The Good School Toolkit' had short-term effects in reducing the

[4]Famous examples of whole-school interventions include the Olweus Bully Prevention Program and the KiVa Anti-bullying Program, created in Norway and Finland, respectively.

likelihood of physical violence from staff to students, as well as the likelihood of absenteeism. In China, an intervention aiming to foster empathy among students by coaching their parents was found to reduce bullying incidents (Cunha et al. 2023). In Pakistan, a student-only intervention that used play and sports to foster adolescents' life skills (e.g., communication and conflict resolution skills, empathy, and cooperation) reduced peer violence and improved mental health (Karmaliani et al. 2020). In Peru, Gutierrez et al. (2018) implemented a randomized control trial in 66 urban schools and studied the effects of a student-only intervention that provided information about the negative consequences of bullying, the importance of standing against bullying, and the available reporting platforms. Similar to my findings, the authors found their intervention increased the willingness to report cases of violence and reduced the likelihood of school mobility.

This paper contributes by analysing the impacts of a nationwide government intervention that followed a staff-only approach. Training school staff is a fundamental first step as many school heads and teachers lack the knowledge and skills to prevent and manage school violence. But evidence on this is mixed. In Jamaica, Baker-Henningham et al. (2019) found that training teachers on positive discipline strategies led to a reduction in the perpetration of violence by teachers, while in Tanzania, a similar intervention implemented in refugee camps had no effects on violence levels (Fabbri et al. 2021). I add to the recent research by showing that a large-scale intervention that trained school leaders not only on positive discipline strategies but also on how to manage and respond to all forms of school-related violence provided a reporting outlet that positively impacted students' desire to remain in the same school. Moreover, even though whole-school interventions are considered ideal (Lee et al. 2015), governments may lack the resources to implement large-scale interventions that provide training and support to both the school staff and the students. Therefore, it is essential to understand the relative impact of alternative interventions. This paper adds to this discussion by showing that interventions focused solely on strengthening the school heads' violence management skills can have similar effects to small-scale student-only interventions (such as the Gutierrez et al. (2018) intervention).

This paper also relates to the literature on human capital formation. The findings add to the few papers exploring the educational effects of school anti-violence or anti-bullying strategies. Similar to student-only interventions (Gutierrez et al. 2018), I provide evidence that the TA - through standalone training to the school heads - influenced the students' decision to switch schools. With regard to learning, the evidence is mixed. Similar to Devries et al. (2015) that studied the effects of a whole-school type of intervention in Uganda, the Peruvian TA did not have short-term effects on learning. On the other hand, Gutierrez et al. (2018) showed improvements in math and language test scores in the medium-term. Further research is needed to understand the short-, medium-, and long-term relationship between school anti-violence strategies and learning outcomes.

Finally, this paper speaks to the literature on school management. Better school management has been found to be positively correlated with educational outcomes (Bloom et al. 2015; Leaver et al. 2019). Yet, there is mixed evidence on the impact of interventions targeting management in public schools. While Romero et al. (2021) and Muralidharan and Singh (2020) do not find that interventions fostering better school management improved educational outcomes in Mexico and India, respectively, Fryer (2017) shows that increasing the principal's management skills led to higher student test scores in the United States. Even though the TA did not address overall school management, the TA trained the school heads on the management of school violence and contributed both to increasing the likelihood of reporting and reducing school mobility, generating supportive evidence about the effects of investing in managerial skills.

The remainder of this paper is organized as follows. In the next section, I describe the administrative and primary data used for the analysis. In section 3, I start by documenting the prevalence of school violence in the country and its link with educational indicators and then explain the institutional background and design of the technical assistance. In section 4, I explain the empirical strategy. In section 5, I analyse and discuss the findings. In section 6, I present different robustness checks. In section 7, I conclude and discuss policy implications and future avenues of research.

2 The Data

2.1 Administrative Data

I construct a panel dataset of 24,211 public schools and 4.6 million students, representing the universe of public schools that were operating throughout 2014 and 2019.[5] The dataset combines five sources of administrative data:

School Census: school level data reported by each school about the school inputs and characteristics (e.g., infrastructure and access to services), school staff characteristics (e.g., number of school staff by type of contract or position, by gender, educational background) and the number of enrolled students (by sex, grade, educational level). The data also includes the latitude and longitude coordinates of each school.

Student Census: student level data of all enrolled students. The dataset has information about the student characteristics (age, sex and education level of the parents) and allows me to construct the educational history of each student, allowing me to identify the students that left school before completing their studies (dropout), as well as the students that move or switch to another school (student mobility).

SíSeVE Reports: report level data that allows me to identify the number of violence reports per school by form of violence, as well as, the age and sex of the victim and the type of perpetrator. The dataset also has information about who registered the incident of violence (e.g., the victim, a confidant of the victim or a witness).

[5]For 98 percent of the schools, I have data over the 6 years. For the remaining 2 percent, schools were created after 2014 but before 2019, so data is only available since the year the school was created. I exclude the schools that were closed before 2019 or created in 2019, as I need data before 2019 to run several robustness checks and data of 2019 to estimate the outcomes of interest. Moreover, the dataset only includes primary and secondary schools with single-grade teaching. It does not include schools that offer primary education in the form of multigrade teaching (9,943 public schools).

Targeted or Beneficiary Schools: school level dataset that indicates the schools that were assigned to receive the intervention.

Evaluación Censal de Estudiantes (ECE): the ECE is a national standardized test on students' knowledge of math and language. Between 2015 and 2019[6], the test was administered to students aged between 13 and 14 years of age, enrolled in second grade of secondary school.[7] In 2018 and 2019, the ECE also included a set of questions to measure the students' perceptions of school violence.

Using a unique identifier by school, I linked all the datasets and construct one dataset at the school level and one at the student level. I mainly use the data from 2019 to analyze the impact of the technical assistance and data before 2019 to assess the validity of the empirical strategy. The main outcomes analysed in this paper include:

- *Likelihood of reporting of violence*: dummy variable that takes the value of 1 if at least one event of school violence was reported, and 0 otherwise.
- *Number of reports of violence*: sum of the reports of violence per school, including reports of any form of violence: physical, psychological, or sexual.
- *Student dropout*: I create an indicator at the student level and the school level. The indicator at the student level is a dummy variable that takes the value of 1 if the student enrols in the academic year t, but does not enrol in the academic year $t + 1$, leaving the school before completing his/her studies. Taking into account that in Peru the academic year starts in March and finishes in December, a student drops out if, for example, he/she enrols in the 2019 academic year, but leaves school before completing his/her studies and does not enrol in school in 2020. Using the student level indicator of dropout, I also construct the school's annual rate of dropout, which measures the proportion of students who drop out in a single year without completing their studies.
- *Student mobility*: I create an indicator at the student level and at the school level. The indicator at the student level is a dummy variable that takes the value of 1 if the student is enrolled at one school during the academic year t, and enrols in a different school for the academic year $t + 1$. I do not consider the structural moves that are required when a student needs to transition to another school because their current school does not offer the educational level they need to enrol on. In Perú this is common for transitions between primary and secondary school. Moreover, the indicator does not consider moves that occur due to school closure. This situation is less common: between 2014 and 2019, 3 percent of public schools closed.[8] Considering this, as defined by Welsh (2017), the indicator can be viewed as an indicator of non-structural mobility. That is, moves that occur when the student could have, in theory, stayed at their previous

[6] In 2017, the ECE was not administered due to El Niño phenomenon that hit the country during that year and generated disruptions in the school year. In 2020, the ECE was not administered due to the COVID outbreak.

[7] The ECE is only administered in schools with more than five students.

[8] Two-thirds of these schools closed before 2019 and the majority are primary schools with multi-grade teaching.

school. Using data about the location of the schools, I create a proxy indicator for residential mobility[9], to differentiate non-school related moves - potentially related to family residential mobility - from school related moves - motivated by student experiences in the school. Finally, using the student level indicator of mobility, I construct the school's annual rate of mobility, which measures the proportion of students who move to a new school in the subsequent academic year.

- *Student Test Scores*: I use the math and language standardized test-scores from the ECE.

Moreover, I create a variety of control variables related to school infrastructure, access to services and characteristics of the school staff, as well as student characteristics, including their age, sex, and parent's level of education (see Appendix B for details).

2.2 Primary Data

I complement the administrative data with in-depth interviews executed with officials at the Ministry of Education and facilitators located at the Local Educational Offices of the Ministry, and with an online survey administered to the school principals from the beneficiary schools. The survey was responded to by 54 percent of secondary schools and 29 percent of primary schools.[10] The primary data allowed me to complement the analysis with information about the program context, design and implementation process, as well as with statistics on the school principal's response to the intervention.

3 Background and Policy Context

3.1 Institutional Background

The Peruvian education system has a decentralized structure with four levels of administration: The Ministry of Education (MINEDU), the Regional Educational Offices (REOs), the Local Educational Offices (LEMOs) and around 24,000 primary and secondary public schools.[11] All schools have the duty to protect their pupils and provide them with a safe environment, free from the harmful effects of violence. However, it was not until 2014 that new legislation and strategies directly targeting school violence were enacted. In 2014, MINEDU published the first National Strategy Against School Violence and formally introduced an online platform to report events of school violence[12] – called SíSeVE.[13]

[9] Even though administrative data does not allow me to observe the address of the student, it has information to identify if the student switched to a school located in the same district, a different district or a different province. Moves within the district and across districts in the same province do not involve, necessarily, residential mobility. However, switching to a school located in a different province requires residential mobility as otherwise, it would be impossible to commute to school. Assuming that moves to a school located in a different province are a proxy for family residential moves, I create an indicator of residential mobility and then create a more precise mobility variable that excludes non-structural moves related to residential mobility.

[10] Overall, the respondents were more likely to work in schools located in urban areas that offered secondary education. See Appendix C for details.

[11] The universe of 24,000 schools correspond to public schools with single-grade teaching. Primary levels cover 6 years of education from age 6 to 11, while secondary levels, cover 5 years from age 12 to 16. 81 percent offer primary education and 40 percent offer secondary education, all following single-grade teaching.

[12] The platform was available since September 2013 but formally introduced it in 2014 as part of the first National Strategy Against School Violence (N°364-2014-MINEDU)

[13] SíSeVE translated into English would be *Yes We See It.*

Despite these efforts, civil servants from MINEDU explained that by 2018 there was an important knowledge gap among the school staff regarding skills and strategies to prevent and manage school violence. Motivated by this, MINEDU designed a technical assistance program that provided training to school principals on the identification and management of violence, the design and implementation of school coexistence rules, and strategies to move from punitive discipline towards positive discipline. This paper focuses on assessing the impact of this training, as it is the first intervention of its kind. Before going into details about the intervention (section 3.3), the next section discusses the phenomenon of school violence in the country.

3.2 Descriptive Statistics

3.2.1 School-related Violence

To date, neither researchers nor practitioners have agreed on a unique definition of school violence. The definition and analysis of the scope of this phenomenon has been limited to a great extent by the survey questionnaires and the data availability (Richardson and Fen Hiu 2018; UNESCO 2019). In this paper, school violence is defined in a broad sense as any behaviour that jeopardizes the intent of the school to be a safe space, free of aggression (Miller and Kraus 2008). It includes different forms of violence - physical, sexual, and psychological - that can emerge as a set of isolated events against different victims or as repeated attacks against the same victim (the latter is known as *bullying*).

School violence survey data suggest that Peru is among the countries with the highest percentage of students between 11 to 15 years of age reporting having experienced school violence. Both data from the 2010 Global School-based Student Health Surveys (GSHS)[14] and the 2013 Third Regional Comparative and Explanatory Study (TERCE)[15] show that 47 percent of the surveyed students said they had been victims of school violence in the last month, levels of violence that are 7 percentage points above the Latin-America average.

Moreover, administrative data sources available in the country provide two important stylized facts. First, **the reported events of school violence have been increasing over time, with bigger jumps recorded after 2018, when the TA was implemented.**

Since 2014, victims or witnesses of violence can report cases of violence through the SíSeVE platform. In addition to SíSeVe, the victims of school violence can report in person at their nearest LEMO or, since 2019, by phone.[16]

The data based on SíSeVE reports has allowed, for the first time in Peru, to detect events of school violence.[17] In the last 6 years, half of the reports were related to incidents of physical violence,

[14]Data on 2882 students aged 13 to 15 years of age. The indicator is constructed using a series of questions in which students indicate whether they have been victims of different forms of violence one or more days during the last month.

[15]Data on 4403 students aged 11 to 12 years of age. The indicator is constructed using a series of questions in which students indicate whether they have been victims of different forms of violence in the last month and whether they fear other students in the school.

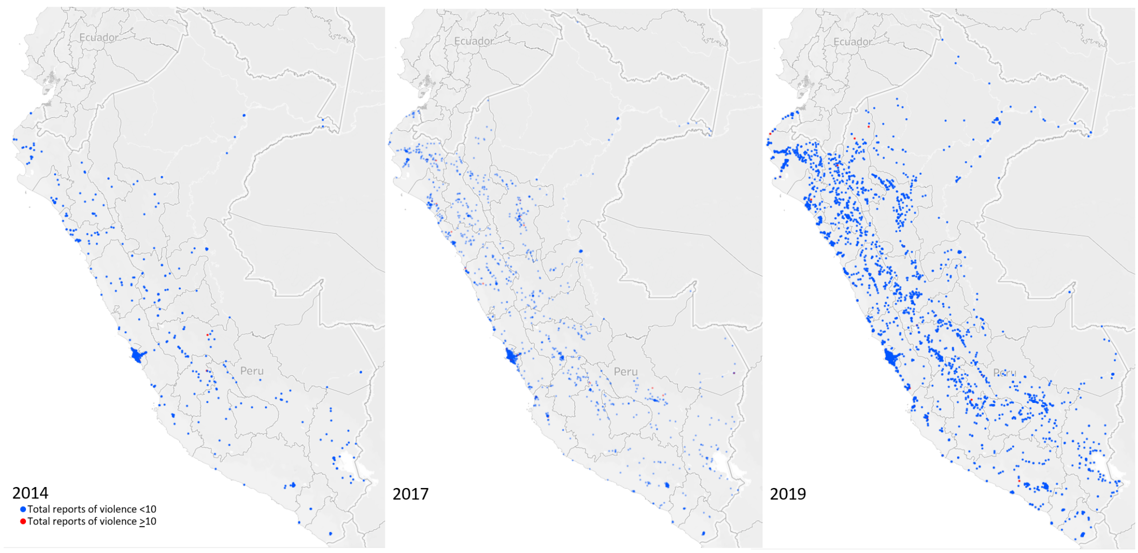
[16]In these latter cases, all the reports are then registered at the SíSeVE platform by government officials to systematize and monitor all cases of violence using a unique platform.

[17]The platform has some distinctive advantages. First, relative to other reporting mechanisms, students might have less fear to report as they do not need to report directly to an adult within the school. Second, witnesses of violence or relatives of the victim can also report any case of violence anonymously, allowing to raise awareness of cases that otherwise would not be identified. Making this option available seems to be particularly important in Peru, as 30

follow up by reports of psychological (32 percent) and sexual violence (17 percent). Consistent with the literature, across all forms of violence, cases were more common in secondary schools, when the victims were 12 to 16 years of age. The most frequent form of violence against girls was psychological violence (37 percent of reports) follow up by 35 percent of reports of physical violence and 29 percent of reports of sexual violence (includes rape, sexual assault and sexual harassment). Among boys, sexual violence is the least common form of violence, with around 5 percent of reports, while two thirds of the reports refer to physical violence and one fourth to psychological violence.

Between 2014 and 2019, 28 percent of public schools registered at least 1 case of violence. Over this period, the number of reports of violence has increased, with the biggest jumps registered from 2014 to 2015, when the SíSeVe Platform was created, and in the period between 2018 and 2019, when the TA was implemented. For instance, in 2017 there were around 90 reports of violence per 100 thousand enrolled students, while in 2019, the number doubled: 207 reports per 100 thousand students. As it can be seen in the maps of Figure 1, the increase in the number of reports of violence is, in part, explained, by the increase in the number of schools reporting cases of violence. Among the schools that registered cases of violence in 2019, 37 percent registered cases for the first time that year.

Figure 1: Schools with reported cases of violence in 2014, 2017 and 2019



Second, **student survey data suggests that not all events of violence are reported**. Among the schools that registered cases of violence, the number of recorded reports may underestimate the true number of incidents of violence. Among the schools that did not register any case, it is uncertain whether incidents of violence occurred. Survey data collected by MINEDU in 2018 and 2019 from students aged between 13 and 14 years old allowed me to explore this further.[18] The survey asked students[19] if they had observed or witnessed incidents of violence perpetrated by other students or the

percent of the cases of violence in the last 6 years, were reported by family members and 50 percent by members of the school staff.

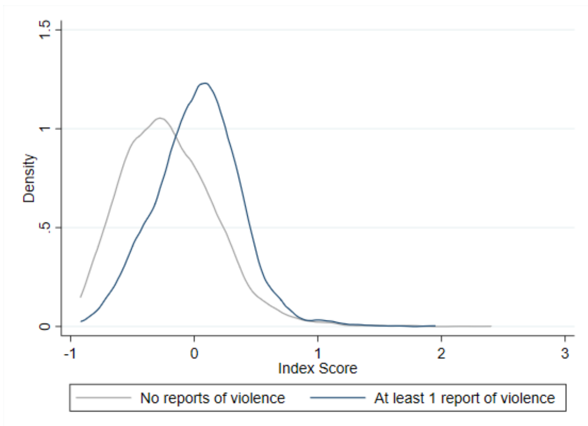
[18] The survey was collected at the end of the academic year, the same day that the National Assessment of Students was administered. It includes students from all secondary schools enrolled in second grade, except for those students enrolled in schools that have less than 5 students

[19] The survey included 6 statements that asked about violence between students and 5 statements that asked about violence teacher to student. The statements did not include questions related to sexual violence. In Appendix F, Figure F.1 summarizes a few of the statements that were included in the survey

teachers. Even though the data does not ask directly if the student was a victim of school violence, it provides an idea of the presence of violence in the school.

The survey data suggests that in all the schools, but to a different degree, students have witnessed incidents of school violence.[20] In 2019, 50 percent of surveyed students said they witnessed at least 1 event of physical and psychological violence between students, and 22 percent witnessed events of both physical and psychological violence from teachers to students.[21] Moreover, using this data, I create an index of perceived school violence and plot the distribution of the index for the schools that registered and did not register incidents of violence in the SíSeVe platform (Figure 2). In both groups of schools, I observe that students witness cases of violence. However, among the schools that registered incidents of violence, the distribution of the index of perceived school violence is shifted towards the right, indicating a higher perception of school violence in these schools relative to the schools where no cases of violence were registered. Even though both measures of violence have to be used with caution due to issues related to underreporting and under coverage, both signal the prevalence of school violence in the country.

Figure 2: Index of perceptions of school violence by school reports of violence



Notes: Grey line shows the distribution of the index of perceptions of school violence for the schools that did not registered any report of violence, while the blue line shows the distributions for the school that registered at least 1 report of violence.

3.2.2 Violence and Educational Outcomes

The prevalence of violence within the schools has proven to have negative effects on educational outcomes - such as learning, school attendance and dropout - for bystanders, victims, and perpetrators of violence (see section 1). Using administrative data from public schools, I explore the rates of student dropout and student mobility and their association with the prevalence of school violence in Peru. In the last decade, the rate of student dropout in public schools has remained stable, around 3 to 4 percent. Regarding student mobility, the rate of students that switched or moved[22] to new schools

[20] For this analysis, I restrict the sample to schools that did not benefit from the intervention. I do this to isolate the potential effects of the intervention from the analysis.

[21] This statistic is estimated by creating a dummy variable that takes the value of 1 if the student said he/she witnessed at least 1 of the statements used in the survey to identify the presence of physical and psychological violence

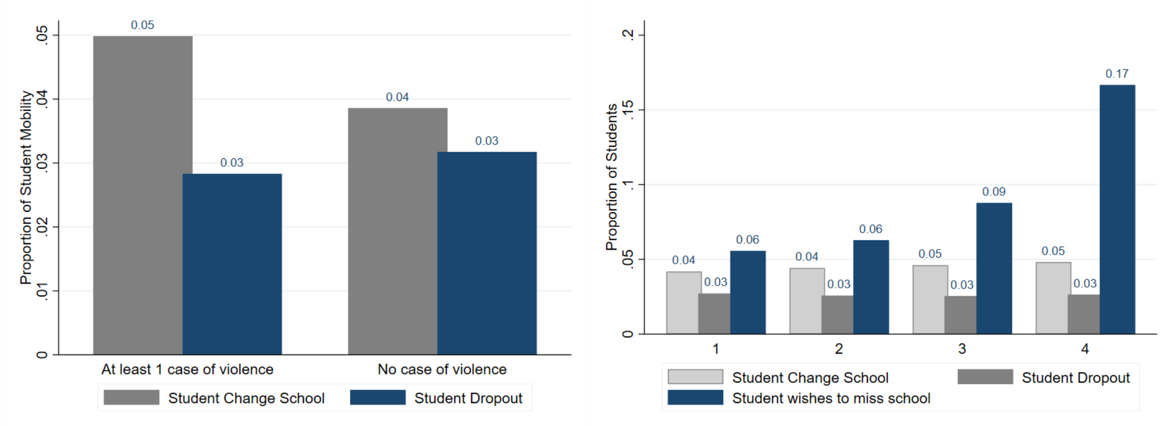
[22] As explained in section 2 , the indicator does not include structural moves.

was between 5 and 6 percent. In the period 2019-20, for every 100 thousand enrolled students, 6 thousand students switched schools

One common source of school mobility is related to residential mobility. The literature considers this as family-specific changes that are not directly linked to a student's experience at school, but instead are linked to family circumstances such as divorce or carer's new job (Rumberger and Larson 1998; Welsh 2017; Burdick-Will et al. 2021). In 2019, for instance, 37 percent and 42 percent of student mobility in primary and secondary schools seemed to be related to residential mobility. The remaining percent of school moves are linked to other factors, including academic preferences and exposure to school violence (Akiba, 2008; Carson et al, 2013; Burdick-Will et al, 2020).

Using the data on reports of violence, I observe that the rate of student mobility is slightly higher in schools that registered at least one case of violence (Figure 3a).[23] I also use the Index of Perceived School Violence to explore the correlation with dropout and student mobility, as well as a proxy for student likelihood of absenteeism.[24] Figure 3b shows the index of perceived school violence disaggregated by quartiles, where the highest quartile indicates the highest levels of perceived school violence. The data suggests that in the schools with higher values in the index of perception of school violence, the rate of student mobility is around one percentage point larger. The proportion of students that would prefer to miss school is also higher among the schools in the fourth quartile related to the remaining schools. In line with the literature, this descriptive evidence suggests an association between violence and education related indicators and motivates exploring the effects of the intervention both on reporting behaviour and educational indicators.

Figure 3: School Violence and Education Indicators



(a) School reported violence

(b) Perceived school violence

Notes: The figure at the left shows the proportion of school dropout and mobility among schools with and without reports of violence. The figure at the right shows the proportion of school dropout and mobility and the proportion of students that wish to miss school by quartile of the index of perceptions of school violence, where higher values indicate worse perceptions of violence.

[23]For this analysis, I restrict the sample to schools that did not benefit from the intervention. I do this to isolate the potential effects of the intervention from the analysis.

[24]This indicator consists of a dummy variable that takes the value of 1 if the student agrees with the following statement 'I prefer to not to attend school'.

3.3 The Technical Assistance

MINEDU designed the technical assistance with the aim of improving the prevention and management of school violence.[25] In each LEMO, a civil servant (from now onwards, LEMO Facilitator) was responsible for implementing the intervention. The intervention was structured in three cycles involving three training sessions, three visits, and three group learning sessions (see Appendix D). Each training session lasted around 4 hours and introduced a new topic in the following order: identification and management of school-related violence (including response protocols by type of violence); positive discipline strategies; and, design and implementation of school coexistence norms.

The training sessions were given at the LEMO or at an alternative venue, while the visits occurred at each school. During each visit, the LEMO Facilitator went to the school to review the concepts discussed during training and to solve any doubts or concerns from the school staff. Moreover, the intervention included group learning sessions. These were designed with the aim of creating a network through which the targeted schools discussed and learnt from each other experiences about managing school violence.

The technical assistance was the first nationwide intervention targeting the topic of school violence directly, and as such, MINEDU prioritized strengthening the capacities of the school principals and teacher representatives[26] that, by law, were responsible for leading the actions towards identifying, preventing, and managing school violence. The decision to focus on the school heads was also motivated by budget constraints that generated a trade-off between reaching more schools versus reaching fewer schools but providing the training to all the school staff.

The TA was implemented across all the LEMO[27] in the country to 2655 schools. To select the beneficiary schools, MINEDU categorized schools into two groups: nucleo schools and adjacent schools. In each LEMO, they targeted 3 nucleo schools and 9 adjacent schools.

Nucleo schools were selected based on the prevalence of violence, the number of enrolled students and their distance to the LEMO. After selecting 3 nucleo schools per LEMO, MINEDU selected the adjacent schools. In each LEMO, the Ministry selected the 3 schools located closest in distance to each nucleo school, targeting in total 9 adjacent schools per LEMO. Even though the distance to the nucleo schools was the main criterion, the number of enrolled students was also part of the selection criteria. When schools were at a similar distance to the nucleo school or when those schools close to the LEMO had few students, MINEDU prioritized the school that had a larger population of enrolled pupils. In each LEMO, the combination between the distance and population criteria had a different degree of importance depending on the dispersion and density of schools.

In the following section, I discuss in detail the empirical strategy that I follow to measure the impact of the 2019 intervention. Considering that the technical assistance was implemented between

[25] This section is based on a set of interviews and conversations held with MINEDU officials throughout 2019 and 2020.

[26] In each public school, the school staff chooses a teacher representative that will support the school principal in all activities related to the management of school coexistence, including school violence.

[27] The system of education only has 220 LEMO, yet in the Region of Callao, the REO was responsible for this as this location does not have a LEMO. Therefore, for the purposes of this study and for simplicity, I refer to 221 LEMO: 220 LEMO and 1 REO.

May and October of 2019 and the outcomes were measured 1 to 4 months after the completion of the intervention (Figure F.2), I will be able to measure the short-term impacts of the intervention on violence-related outcomes (*number of reports of violence*) and education-related outcomes (*school dropout, school mobility, and learning*).

4 Estimation Framework

Like most public large-scale interventions, the beneficiary schools of the TA were not randomly assigned to the intervention. Therefore, I will exploit the eligibility rules to find a valid group of schools that was not assigned to receive the intervention and that is unlikely to differ from the beneficiary schools in terms of their observable and unobservable characteristics.

Eligibility rules: Given budget constraints, in each LEMO, MINEDU only chose 12 beneficiary schools: 3 nucleo schools and 9 adjacent schools (see section 3.3). I will focus the analysis on the adjacent schools as the eligibility rules create an exogenous variation that allows me to estimate the impact of the TA in these schools. Another important reason to focus on the adjacent schools is that 90 percent of nucleo schools were targeted to receive another intervention at the end of the school year, making it harder to disentangle the effect of the TA for these schools.

The two variables used to select adjacent schools were the distance to the nucleo schools and the number of enrolled students. MINEDU mapped all the public schools in the country, and, for each LEMO, selected the 9 schools that were closer in distance to one of the 3 nucleo school[28] and that had the highest number of enrolled students. Even though MINEDU did not officially create a ranking, they explained that the selection process mimicked a ranking procedure under which the top 9 schools in each LEMO were assigned to receive the intervention.

The ranking procedure and the top-9 threshold rule provide an opportunity to analyse the impacts of the intervention using a regression discontinuity design. This method, introduced by Thistlethwaite and Campbell (1960), allows to analyse the impact of an intervention when the assignment to treatment is determined by an assignment or running variable that exceeds a known cut-off-point. In the context of the TA, I will exploit the fact that the eligibility criterion mimics a ranking procedure under which the schools in the top-9 of the ranking were assigned to receive the intervention and those just above the threshold rule were not chosen because they were a few kilometres further away from the nucleo schools and/or because they had a lower number of enrolled students.

MINEDU only kept a record of the schools that were targeted to receive the TA. Therefore, to study the potential exogenous variation generated by the eligibility criteria, I estimate the ranking of schools as follows (see Appendix E for a detailed explanation):

1. *Create a 'distance to nucleo' ranking*: I estimate the distance in kilometres between all the public schools and the nucleo schools within a LEMO. Then, I rank the schools within each LEMO in ascending order based on their distance to each nucleo school.

[28]This meant selecting 3 adjacent schools per nucleo school.

2. *Create a 'population' ranking*: I rank the schools within each LEMO in descending order based on the number of enrolled students (from now on I will refer to this as the population ranking), where schools ranked first, represent the schools that had a larger number of enrolled students.

3. *Create a score per school based on the distance and population ranking*: Qualitative interviews with MINEDU revealed that the importance given to the distance and population variable varied by LEMO, mainly depending on the density and dispersion of schools. Therefore, I explore 23 different weighting schemes or combinations of weights, where the weights assigned to the distance and population variable range between 0 to 1 (see Table E.1 for the full list of weights). This means that for each school I create a score for each of the 23 weighting schemes following equation 4.1.

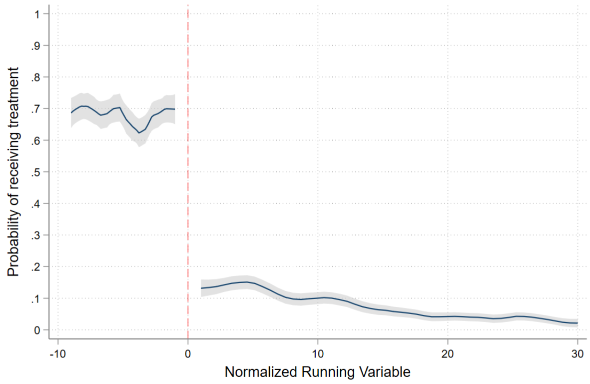
$$Score_{ijw} = RankDistance_{ij} W^{w}_{distance} + RankPopulation_{ij} W^{w}_{population},$$
$$\text{where i=1 to N school, j=1 to 221 LEMO and w=1 to 23 weighting schemes} \tag{4.1}$$

4. *Create the ranking based on the score obtained in the previous step*: I then rank schools in ascending order based on the score obtained after estimating equation 4.1 for each weighting scheme. For each LEMO, I use the weighting scheme that yields the highest predictability rate.[29] Finally, I normalize the chosen ranking to zero and this becomes my assignment or running variable.

Figure 4 shows the relationship between the treatment status defined by MINEDU (dummy that takes the value of 1 if the school was assigned to treatment, and zero otherwise) and the running variable. It clearly indicates the presence of discontinuities in the probability of treatment and reveals that the probability jumps by less than one, suggesting that a fuzzy RDD can be a promising empirical strategy. Interviews with MINEDU, as well as primary survey data collected from the LEMO facilitators, indicate that in a few cases, some exceptions were made. Even though the selection of the beneficiary schools was done by the central office of MINEDU, the LEMO could suggest modifications. As a result, in a few LEMO 1 or 2 exceptions were made. The main reasons for this were related to logistic concerns (for instance, to prioritize a school that was also closer to the LEMO office) or to prioritize schools that could be located further away but were considered to be more vulnerable in terms of school violence. Considering that the LEMO suggestions had to be approved by MINEDU and that the LEMO had to provide valid and verifiable reasons, it is unlikely that favouritism towards specific schools influenced the selection criteria. This is also unlikely considering that two-thirds of the LEMO Facilitators were hired for the first time in 2019. Moreover, considering that MINEDU used a map of schools to inspect visually which schools were closer to the nucleo schools for each of the 221 LEMO, it is likely that random human error also explains the fact that some schools above the cut-off were treated.

[29] In other words, for each LEMO, I use the weighting scheme (or combination of weights) that predicts more closely the Ministry of Education's official selection of beneficiary schools. I do this by comparing the number of eligible schools predicted by my algorithm, and the number of schools selected to be treated by the Ministry of Education.

Figure 4: Probability of Treatment



Note: The dashed vertical red line represents threshold cut-off. The solid line represents the relationship between the treatment status defined by MINEDU and the running variable. The gray area show the 95% confidence intervals.

Fuzzy RDD: I use a standard two-stage least square (2SLS) procedure to estimate the program impacts. In the first stage all the coefficients of the equation 4.2 are estimated using a linear probability model, where D_{ij} is a treatment status dummy that takes the value of 1 if the school i located in the province where the LEMO j operates was assigned to treatment, and 0 otherwise. T_{ij} is an eligibility dummy that takes the value of 1 for those schools located below the threshold, and 0 otherwise.[30] $ranking_j$ corresponds to the running variable and λ_j represents LEMO fixed effects.

$$D_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 ranking_{ij} + \lambda_j + \mu_{ij} \tag{4.2}$$

In the second stage, I estimate the following specification[31]:

$$y_{ij} = \alpha_0 + \alpha_1 \hat{D_{ij}} + \alpha_2 ranking_{ij} + \gamma_j + \epsilon_{ij} \tag{4.3}$$

where y_{ij} represents the outcome variable of interest for school i located in the province where the LEMO j operates. $\hat{D_{ij}}$ represents the predicted probability of the treatment status dummy. The instrumental variable estimates of equation 4.3 use the discontinuities in the relationship between the treatment status and the eligibility dummy to identify the causal effect of the intervention for the adjacent schools, where α_1 is the coefficient of interest that shows the local average treatment effects (LATE). γ_j represents LEMO fixed effects.[32] I also estimate another specification that includes an interaction between the running variable and the treatment dummy (instrumenting with an interaction term between the running variable and the eligibility dummy). Section 6 also discusses the various

[30]The cut-off rule is 9, however, considering I normalized the running variable to zero, $T_{ij} = 1$ when $ranking_j \leq 0$.

[31]For individual level outcomes, equation 4.3 would be: $y_{sij} = \alpha_0 + \alpha_1 \hat{D_{ij}} + \alpha_2 ranking_{ij} + \gamma_j + \epsilon_{sij}$, where s refers to the student,i to the school and j to the LEMO.

[32]The addition of fixed effects allows me to study within LEMO variation. Section 6 shows results remain similar if we remove the fixed effects.

smooth forms of the running variable.

The running variable used in this paper is discrete. This is common in other RDD applications that use, to mention a few, age (Lalive et al. 2006; Lemieux and Milligan 2008), date of birth (Card and Shore-Sheppard 2004; Oreopoulos 2006; McCrary and Royer 2011) and number of employees (Hahn et al. 2001) as their assignment variable. Discrete running variables do not introduce particular complications for the parametric estimation (Lee and Lemieux 2010). As explained by Lee and Card (2008), if the discrete variable only takes a few values and the gap between the closest value and the threshold is high, there could be few observations just above and below the threshold and the econometrician might need to move away from the threshold, and hence, has to impose a functional form. This is also common practice when using continuous running variables and therefore, it is suggested to analyse if results are robust to changes in the functional form. I do this and observe that results remain consistent both when using a linear and a quadratic functional form at different windows of analysis (see section 6).

Following Kolesár and Rothe (2018), I use heteroskedasticity-robust standard errors clustered by LEMO. An alternative method, suggested by Lee and Card (2008) and used frequently in empirical work (Oreopoulos,2006; Urquiola and Verhoogen, 2009; among others) involves clustering the standard errors by the running variable. However, Kolesár and Rothe (2018) find that this approach has poor coverage properties. The authors find that clustering by the running variable provides inappropriate narrow confidence intervals and suggest using more conservative heteroskedasticity-robust standard errors. In Appendix G.5, I show the estimates following both procedures and observe that results remain similar and that, as expected, standard errors are smaller when clustering by the running variable.

The discrete running variable allows for 9 different windows of analysis. The main results presented in the paper use a window of data around the discontinuity of ±5, a neighbourhood that contains 2,193 schools. I chose this window since it represents a middle point, without being too close or too far from the cutoff. Using a data-driven approach developed by Calonico et al. (2014)[33], I also explore the optimal bandwidth choice and confirm that the optimal bandwidths are between 4 and 6, depending on the type of outcome and the order of the polynomial. As a robustness check, I also run the analysis in all 9 windows of analysis (see section 6).

Validity of the fuzzy RDD: For the Fuzzy RDD to be a valid empirical strategy there has to be imprecise control over the running variable. The central office of MINEDU, located in Lima, selected the beneficiary schools based on the eligibility criteria. The schools had no prior knowledge about the criteria and, even if they did, it is unlikely that they could have manipulated the variables. First, the distance variable is based on the longitude-latitude coordinates of each school to the Nucleo schools. Schools have no control over their latitude-longitude coordinates and all, but 2 treated schools, were created prior to the implementation of the intervention in 2019, making implausible the prospect of creating schools in a specific location just to benefit from the TA. Second, administrative data on

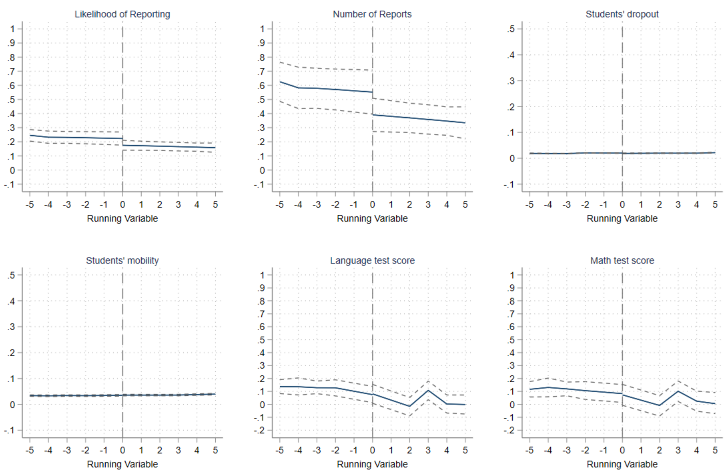
[33] The method uses one common mean squared error (MSE) optimal bandwidth selectors and adjusts for mass points or repeated observation in the running variable that is common in the settings with discrete running variables.

enrollment was registered by the schools prior to the intervention.[34]

The validity of the Fuzzy RDD also relies on showing suggestive evidence that all relevant factors besides the treatment status vary smoothly at the threshold. I explore this formally by estimating equation 4.3 but using as a dependent variable the predetermined outcomes and covariates (Lee and Lemieux 2010) (see Figures 5 and 6, and Appendix F, Table F.2). I do not find discontinuities in the predetermined outcomes and the majority of the baseline covariates. The covariates for which I find a discontinuity include indicators of whether the school staff was chosen by meritocracy and whether the school offers secondary education. Schools in Peru can offer both primary and secondary levels of education or only one level of education. The discontinuity in this variable indicates that below the threshold the proportion of schools with secondary level is higher. The fact that these covariates are 'locally' unbalanced across different windows of analysis might be a source of concern. However, even though these baseline covariate jumps at the threshold, the estimates remain similar after the inclusion of covariates, suggesting that the validity of the Fuzzy RDD is not compromised.

Furthermore, considering that I follow a 2SLS procedure, it is crucial to discuss the relevance of the instrument. In the result tables presented in the section 5, I present both the F-statistic and the p-values of the Anderson-Rubin (AR) test. The first-stage F-statistic is above the numerical threshold of 10 that is discussed by Staiger and Stock (1997) and Stock and Yogo (2005) and that is commonly used in applied work to confirm the relevance of the instrument. Moreover, Stock and Yogo (2005), Andrews et al. (2019) and Lee et al. (2022)[35] indicate that the AR test is a preferred test for the just identified model as it is robust to weak instrumental variables. Therefore, I also report the p-values from the AR test and observe that the reported values support the validity of the instrument.

Figure 5: Continuity test of predetermined outcome variables

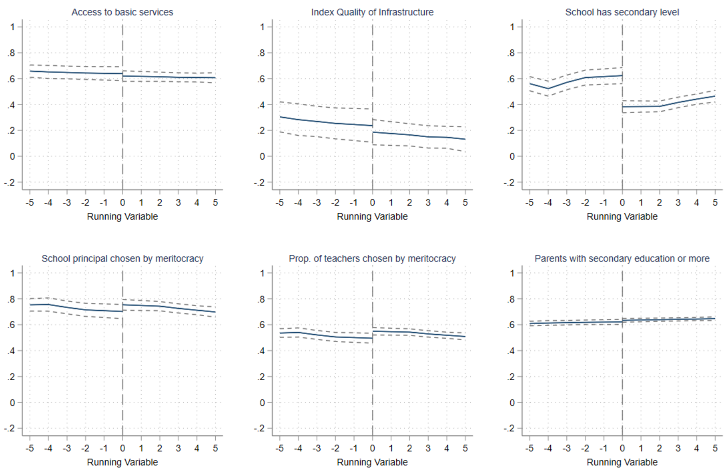


Note: The dashed horizontal lines show the 95% confidence intervals.

[34]It is important to mention that studying potential sorting around the threshold as proposed by McCrary (2008) is not possible considering the evaluation design used in this paper. This is because the ranking only allows each school to take a unique position in the ranking. Considering that there are 221 LEMO, there are only 221 schools located in positions 1, 2, 3 and so on in the ranking. Therefore, by construction, sorting around the threshold cannot be studied as proposed by McCrary (2008) (see Figure F.3 in Appendix F).

[35]Lee et al. (2022) discuss that the threshold value of 10 is not accurate enough to assess the relevance of an instrument. The authors suggest using alternative procedures when the F-stat is below 104.7, including the use of the AR test.

Figure 6: Continuity test of predetermined covariates



Note: The dashed horizontal lines show the 95% confidence intervals.

Exposure to a similar intervention in 2018: in 2018, MINEDU implemented an intervention that had a similar objective relative to the 2019 TA but had differences in terms of the selection criteria of beneficiary schools, the scope, and the degree of implementation fidelity. In 2019 adjacent schools were chosen based on their distance to the nucleo school and the number of enrolled students, while in 2018 the schools were chosen based on the number of reports of violence, the number of enrolled students and the distance of each school to the LEMO. The topics and number of activities also changed over time. In 2019 the TA included group learning sessions and the curricula covered the topic of positive discipline. It is also important to keep in mind that in 2018, the exposure to programme activities was heterogeneous: in 40 (18 percent) LEMO it was not possible to implement the 2018 intervention and in around 10 percent of LEMO fewer activities were implemented due to logistic constraints.[36] Therefore, exposure to 2018 activities can be viewed as a light-touch intervention.

MINEDU did not exclude schools exposed to 2018 intervention from the possibility of receiving treatment again in 2019 if they fulfilled the eligibility criteria. Therefore, I do not drop from the sample the schools that were exposed to 2018 activities to be able to replicate the ranking. I observe that in all the windows of analysis, approximately 30 percent of schools were assigned to treatment both in 2018 and 2019 and 15 percent of schools were assigned to receive treatment only in 2018. Taking this into account, I run a placebo regression in which I use as a dependent variable the treatment status in 2018 and I observe there is no jump at the discontinuity (Figure I.1, Appendix I). This confirms that treatment status in 2018 is independent to the eligibility in 2019 and provides more confidence over the estimates. Yet, there could still be concerns regarding ex-ante differential levels of knowledge about the school management of violence. Taking this into account, in section 6 and Appendix I, I discuss the additional checks that I do to explore the effect of having in the sample a few schools that were exposed to at least one activity of 2018 intervention.

[36]The main difficulty was hiring the facilitator responsible for implementing the intervention, so in several LEMO the activities started around 4 months later than planned

Shortcomings of the empirical strategy: The use of a fuzzy RDD allows me to overcome threats to internal validity, but it also has a few challenges. First, the fuzzy RDD uses only a sample around the threshold. As such, the estimates are less likely to be relevant for the schools located far from the threshold. These are schools that are more likely to be rural and with a smaller number of enrolled students. Second, the statistical power to detect heterogeneity is limited.

5 Results

5.1 Reporting Violence

5.1.1 Likelihood of reporting and number of reports:

I first examine whether the intervention had an effect on the likelihood of reporting incidents of school violence. Table 1 shows the intention to treat effects (ITT)[37] and LATE estimates for all the schools that fall within the window of analysis of ± 5. I observe that among eligible schools, the likelihood of reporting a case of violence increased by 15 percentage points.[38] The estimates are robust to the incorporation of covariates (columns 2 and 5), as well as the inclusion of an interaction between the running variable and the treatment dummy[39] (columns 4 to 6). In columns (3) and (6), I add a dummy variable that takes the value of 1 if the school was exposed to at least 1 activity of the 2018 intervention. Consistent with the fact that eligibility to become a beneficiary of 2019 and 2018 interventions are independent, we observe that results remain almost unchanged after the incorporation of this covariate.

I also analyze the impact of the TA on the number of reports of school violence. Overall, I observe that in the schools below the threshold (treated schools), the number of reports of violence increased, on average, by 1 report.[40] To put this in context, as can be seen in Figure 7, schools mainly report between 1 and 3 reports of violence, with a higher proportion of schools reporting cases of violence among the schools below the threshold. The mean reports of violence in the comparison group was 0.360, an average that is lower than 1 because many schools did not report any case of violence. Conditional on reporting cases of violence, the mean and median number of reports of violence among the schools in the comparison group was 3 and 2, respectively. These figures would suggest that the average increase in 1 report is not trivial. Moreover, the increase in reporting might be capturing both an extensive margin increase that comes from the fact that a higher proportion of schools reported

[37]The ITT are estimated using the following equation $y_{ij} = \delta_0 + \delta_1 T_{ij} + ranking_{ij} + \gamma_j + \epsilon_{ij}$, where T_{ij} is the eligibility dummy that takes the value of 1 for those schools located below the threshold, and 0 otherwise; and δ_1 represents the ITT.

[38]Considering that my outcome variable is a binary variable, as a robustness check I also use a probit model to estimate the ITT and a bivariate probit to estimate the LATE. The coefficient estimates are very similar to those estimated assuming a linear probability model. See Appendix H.

[39]We instrument for this term with an interaction term between the running variable and the eligibility dummy (takes the value of one if a school falls below the threshold).

[40]Number of reports of violence is a non-negative limited dependent variable that is skewed to the right and has many zeros. In such cases, instead of assuming a linear model - as I do for the regressions presented in this paper -, it is suggested to use a non-linear model, particularly an exponential model or Poisson regression model. I also estimate the ITT estimates using a Poisson regression and observe that there is an increase in the reports of violence by more than 100%. A result that is consistent with my findings from the ordinary least squares estimates given that the average number of reports of violence above the threshold is 0.360 reports, so the average increase to 1 report of violence across treatment schools corresponds to an increase that is above 100%.

Table 1: Likelihood of Reporting Violence

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV - LATE	0.223***	0.133*	0.134*	0.232***	0.144**	0.146**
	(0.0638)	(0.0687)	(0.0684)	(0.0647)	(0.0699)	(0.0696)
ITT	0.130***	0.0678*	0.0685*	0.130***	0.0683*	0.0690*
	(0.0403)	(0.0374)	(0.0372)	(0.0403)	(0.0374)	(0.0372)
F-stat	205.2	182.3	182.2	86.18	100.5	101.3
Anderson-Rubin Test P-values	0.00145	0.0710	0.0670	0.00118	0.0800	0.0751
N	2193	2186	2186	2193	2186	2186
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var and Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating equation 4.3 for the window of ±5. The first row shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Columns (4) to (6) include an interaction between the running variable and the treatment dummy (instrumenting for this term with an interaction term between the running variable and the eligibility dummy that takes the value of one if a school falls below the threshold). Columns (2) and (5) include school covariates (i.e., school access to basic services, infrastructure quality, personnel chosen by meritocracy, school offers secondary level of education, proportion of parents with secondary education, and pre-treatment levels of the outcome variable), and columns (3) and (6) incorporate a dummy variables that takes the value of 1 if the school was exposed to 2018 intervention. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

a case of violence for the first time among the treated schools, as well as an intensive margin change explained by increases from reporting at least 1 incident of violence to reporting 2 or more incidents of violence.

The data suggests that the increase in the number of reports of violence is driven mainly by secondary schools (see Table F.3, Appendix F). This result is in line with other empirical research that suggests that children are more likely to be victims of school-based violence between the ages of 11 to 14 which correspond to the first years of secondary schooling (Eslea and Rees 2001; Menesini and Salmivalli 2017; Aboagye et al. 2021).[41]

Moreover, the data allows me to observe the type of incident of violence and who reported it. Regarding the former, I keep the sample of schools that had reports of violence to explore some descriptive statistics. I observe that below and above the threshold the majority of reports of violence were cases of physical violence, followed up by cases of psychological and sexual violence (see Table F.4, Appendix F). Moreover, conditional on reporting, I observe that the schools below the threshold (treated group) had a significantly higher proportion of reported cases of physical violence. To explore this further, I use my main sample and estimate equation 4.3 using as an outcome the different forms of violence.[42] The results confirm that among treated schools, there was a higher number of reported cases of physical violence (see Table F.5, Appendix F). This would suggest that the intervention might have increased the likelihood of reporting forms of violence that are considered more severe relative to psychological forms of violence.

[41]Importantly, the evidence also suggests that the prevalence of school violence by age differs by type of violence, where cases of physical and psychological violence between 8 and 10 years old are also frequent (UNESCO, 2019), and cases of sexual violence are more prevalent in later years (Evans et al. 2023). Yet, data to analyze this with precision is still scarce.

[42]For example, to capture the likelihood of reporting a case of physical, psychological or sexual violence separately, I create three dummy variables. Each dummy variable takes the value of 1 if the school registered a specific form of violence (e.g., physical violence) and zero otherwise (where zero includes no reports of violence, as well as reports on other forms of violence).

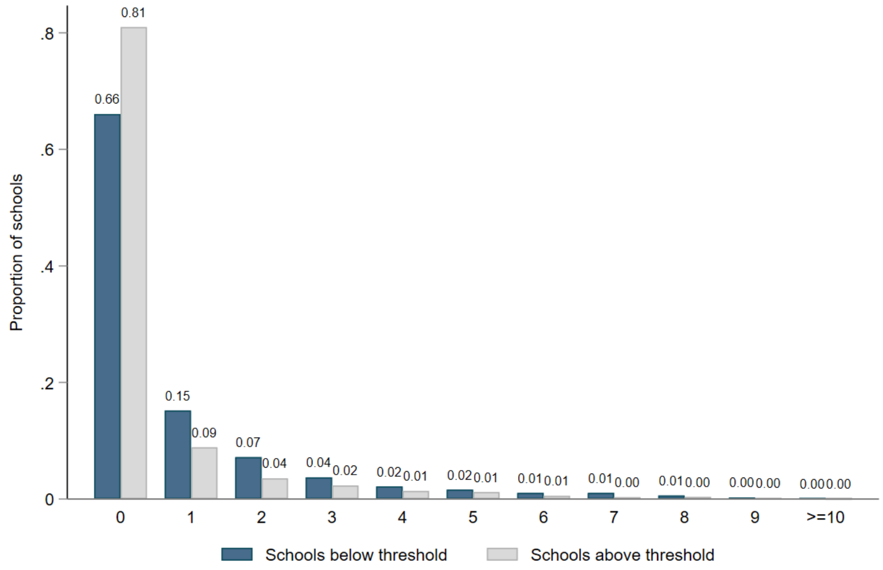
Regarding the agents who reported the incidents of school violence, I observe that both in the treatment and comparison schools, almost all the cases were reported by agents other than the victim, suggesting that the victim generally seeks support from others. The main agents reporting included members of the school staff (mainly, the school principal and the school teachers), and the family of the victim. Conditional on having reports of school violence, I note that there is not a statistically significant difference in the person that reported the incidents of violence between the treatment and comparison schools (see Table F.4, Appendix F). This is an important finding as we could worry about evaluation-driven effects under which the school staff would have registered reports just to show compliance with the intervention.

Table 2: Number of Reports of School Violence

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV - LATE	0.940***	0.679**	0.684**	0.930***	0.663**	0.668**
	(0.270)	(0.290)	(0.289)	(0.259)	(0.278)	(0.277)
ITT	0.547***	0.347**	0.350**	0.546***	0.346**	0.349**
	(0.166)	(0.155)	(0.155)	(0.166)	(0.155)	(0.154)
F-stat	205.2	181.8	181.8	86.18	101.3	102.3
Anderson-Rubin Test P-values	0.00117	0.0267	0.0250	0.00257	0.0810	0.0776
N	2193	2186	2186	2193	2186	2186
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating equation 4.3.The first row shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Columns (4) to (6) include an interaction between the running variable and the treatment dummy (instrumenting for this term with an interaction term between the running variable and the eligibility dummy that takes the value of one if a school falls below the threshold). Columns (2) and (4) include school covariates (i.e., school access to basic services, infrastructure quality, personnel chosen by meritocracy, school offers secondary level of education, proportion of parents with secondary education, and pre-treatment levels of the outcome variable), and columns (3) and (6) incorporate a dummy variable that takes the value of 1 if the school was exposed to 2018 intervention. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 7: Proportion of schools by the number of reports of violence



Notes: The figure shows the proportion of schools by the number of reports of violence below and above the threshold.

5.1.2 Changes in reporting behaviour versus changes in actual violence levels:

One plausible explanation for these results is that they are reflecting a change in the reporting behaviour rather than actual increases in school violence. Reporting violence is a necessary step to identify the prevalence of school violence and to allow the authorities to take action to stop violence from happening again. However, not everyone will be willing to report and the intention to report does not necessarily translate into actual reporting (Tomczyk et al. 2020). An individual who is either a victim, a confidant of the victim or a witness of school violence has to decide between two mutually exclusive actions: to report the incident of violence or to stay silent and not report the incident (considered to be the *status quo*). From a rational economic point of view, the individual will decide whether or not to report depending on the benefits and costs associated with each action and will choose to report if this action yields a higher expected utility relative to the status quo[43].

School, family, and individual factors will determine the weight individuals assign to the benefits and costs[44] of their set of actions and will influence their decision to report violence. The intervention could have shifted reporting decisions mainly through changing school factors, particularly by improving the school's ability to address the issue of school violence. Taken together, the following pieces of evidence support the story that the intervention could have reduced barriers to reporting that led to a greater willingness to report events of violence.

[43]Based on rational choice models (Simon 1955), criminology theory (Becker 1968; Pogarsky et al. 2018) and help-seeking behaviour models (Pescosolido 1992)

[44]The benefits and costs of reporting will be mainly non-pecuniary. The benefits relate to improvements in wellbeing that come from feelings of safety, self-protection, protection of others, and retribution of justice (Skogan 1984). The costs, on the other hand, relate to the opportunity costs of the individual's time (i.e., time spent reporting the case at a police station), subjective costs, and potential external punishments for reporting. Subjective costs mainly exist in the mind of each decision-maker and relate to feelings of embarrassment, shame, and guilt. Potential external punishments are related to fears of retaliation from the offender and disapproval or judgment from peers (Oliver and Candappa 2007; Sulak et al. 2014; Xie and Baumer 2019).

A. School practices and barriers to reporting

Qualitative data and primary-survey data allow me to explore whether the schools changed practices after the intervention. Through in-depth interviews, I learned that before the intervention, not all of the school community knew how and where to report and that there were fears and uncertainties related to reporting. Evidence from an online school survey that I administered to the school heads of the beneficiary schools 4 to 7 months after the intervention, indicates that more than two-thirds of the beneficiary schools worked for the first time on tasks that could have reduced barriers to reporting. This included practices related to the dissemination of information about where to report cases of violence, the creation of spaces for students to report and talk about school violence, the monitoring of cases of school violence, and the execution of general meetings with the school community to talk about school violence.[45]

To explore this further, I use administrative data from a School Census Survey collected by MINEDU in 2020 that is available for all public schools in the country. This self-reported survey was responded to by school principals and contained information about several school practices, including a few practices related to school violence management. The practices included: (i) Designing school coexistence rules, (ii) Appointing a teacher representative responsible for school coexistence, (iii) Registering the school to the online reporting platform to monitor cases of violence, (iv) Making available a school incident logbook, (v) Designing classrooms work-plans that incorporate the topic of school violence, and (vi) Executing violence prevention activities. Despite not being an exhaustive list of practices, this data would provide an indication of schools' actions against school violence.

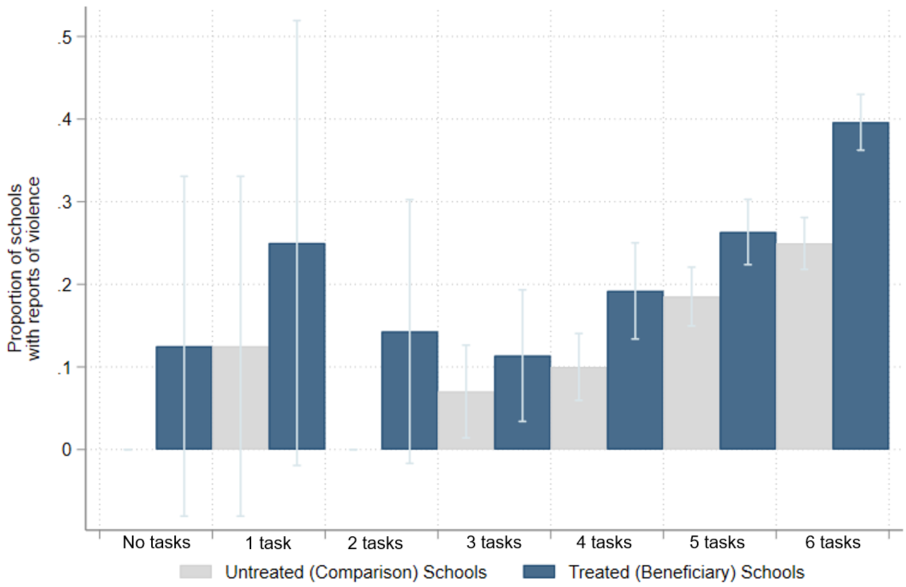
I use this data to create an indicator of the total number of school practices related to the management of school violence and study whether there were differences between the schools below and above the threshold. Appendix F, Table F.6, columns (2) and (4), suggest that treated schools implemented, on average, more practices. The data is collected without differentiating by school level, yet I categorize schools based on whether they offer primary education, secondary education or both levels of education. The estimated coefficients would suggest that results are driven by schools that offer secondary education. In these schools, the treated schools implemented one more practice relative to the comparison schools. This finding provides supporting evidence about changes in school violence management practices that might have influenced the willingness to report cases of school violence. For example, these practices could have reduced information barriers related to where and how to report, as well as potentially increased students' trust in the school's ability to deal with school violence.[46]

[45] The survey is based on self-reported data. We might worry that few school principals provided biased responses to 'look good'. Three things that might reduce these concerns are the following. First, survey participants were aware that I did not work for the Ministry. Second, several survey participants, particularly the ones who responded to the survey on the phone, used the survey as an opportunity to be critical about the intervention and provided feedback about potential improvements for future interventions, giving the impression that they were not primarily interested in giving a 'good impression'. Third, the LEMO Facilitator survey, administered to those in charge of providing the training, also suggests that 82 percent of the schools implemented changes in their school violence management practices.

[46] Unfortunately, I am not able to provide conclusive evidence on which specific tasks matter more. The data mainly allows me to observe that the tasks with the most variation are the ones related to the execution of school violence prevention activities and the incorporation of the topic of school violence in class work-plans. However, I do not observe

To explore the relationship between the number of executed practices and my outcomes of interest, I plot the proportion of schools with reports of school violence by the number of practices executed in the treated and comparison schools. Figure 8 shows that both treated and comparison schools implemented practices. Yet among treated schools that implemented more practices, reporting of violence was higher. This would indicate that in treated schools the intervention might have changed the way the practices were executed, the frequency of implementation of each practice or the number of sub-tasks that treated schools have implemented but that are not captured in my indicator of school practices. My results would overall suggest that changing school violence management practices matter to shift reporting decisions.

Figure 8: Proportion of schools with school violence reports by the number of implemented practices



Notes: The figure shows the proportion of schools with registered cases of school violence by the number of practices executed in the treated and the comparison schools.The number of schools working on less than 3 tasks is small so we observe wider confidence intervals in these cases.

B. Perception of school violence and violence levels

An alternative story could be that the increase in reporting is reflecting an increase in violence levels. It is widely accepted in the literature that reporting data does not necessarily reflect the true prevalence of violence. Empirical research studying different forms of violence, including domestic and sexual violence, has shown that compared to survey-victimization data, report-based data generally underestimates victimization rates (Skogan 1984; García-Moreno 2005; Doleac and Carr 2016; Xie and Baumer 2019). Administrative data, collected by MINEDU from students aged 13 to 14, revealed that around 50 percent of students have witnessed incidents of physical and verbal violence in their schools. Even though this data is noisy, it suggests that not every event of violence is reported. For instance, in more than 90 percent of the schools that reported a case of violence for the first time

statistically significant differences between the treated and comparison schools in the implementation of each individual task. Therefore, it might be the case that working on a set of tasks rather than in isolated tasks is what matters the most.

in 2019, 2 out of 10 students had witnessed cases of physical and psychological violence in 2018.[47] This evidence points to the importance of reducing the barriers to reporting and normalizing the importance of speaking up when facing or witnessing violence at school.

Using this student-level dataset, I create three indexes of perceptions of school violence, where higher values indicate worse perceptions of school violence. The indicators are built using factor analysis based on the student responses to several questions regarding whether they witnessed events of physical and psychological violence in the school and whether this was perpetrated by students or teachers. Table 3 shows the impact of the TA in the index of perception of violence between peers, teacher to student violence and overall perceptions of violence. I do not observe statistically significant differences in the indexes of perception of school violence between the treated and comparison schools. This finding serves as suggestive evidence that there was not an increase in violence levels among treated schools.

In sum, my short-term results are more likely to inform about improvements in reporting rather than an actual increase in violence. This can be interpreted as good news as greater reporting is a necessary first step to identifying and dealing with events of school violence.

Table 3: Index of perception of school violence

	Between Students		Teacher to Student		Overall	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV-LATE	0.0274	0.0213	-0.00855	0.00512	0.0198	0.0195
	(0.0642)	(0.0678)	(0.0521)	(0.0572)	(0.0638)	(0.0681)
F-stat	77.84	24.24	76.63	24.02	77.57	24.55
Anderson-Rubin Test	0.670	0.879	0.871	0.786	0.757	0.953
N	54775	54775	55177	55177	54190	54190
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School & Individual Level Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Exposure to TA 2018 Covariate	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	Yes	No	Yes	No	Yes

Notes: The outcome variables refers to an index of perception of school violence where higher values of the index indicate signal a higher prevalence of violence between students, teacher to students and overall violence. The first row shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Columns (2), (4) and (6) include an interaction between the running variable and the treatment dummy (instrumenting for this term with an interaction term between the running variable and the eligibility dummy that takes the value of one if a school falls below the threshold). Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Staying at School

Substantive research has shown the negative consequences of school dropout (Lleras-Muney 2005; Oreopoulos 2007; Heckman et al. 2011; Gubbels et al. 2019), and, even though it has been less explored, empirical research has also found that school mobility is correlated, in the long run, to student dropout (Rumberger and Larson 1998; Gasper et al. 2012). Moreover, switching schools has been found to be associated with logistic and administrative costs (e.g., providing documentation, new transportation arrangements) as well as psychic costs (i.e., stress or anxiety from adjusting to new environments, making new friends or building new peer networks) that may affect student's

[47] See section 3 for details about this dataset.

psychological well-being, social and academic engagement (Hanushek et al. 2004; Schwartz et al. 2017; Welsh 2017), with some evidence showing poorer student performance among students that switched schools(Hanushek et al. 2004). Considering this and that exposure to school violence is correlated with both higher student dropout and student mobility (see section 1), I use student level data to explore the impact of the TA on these outcomes.

The estimated coefficients suggest that the intervention did not have statistically significant impacts on the likelihood of student dropout (Table 4, Panel A). However, it reduced the likelihood of student mobility. The outcome of student mobility refers to non-structural school moves that occur when the student could have, in theory, stayed in their previous school. In these cases, switching schools could have been motivated by family-related factors or school-related factors. The former refers mainly to cases of family residential mobility - due to the carer's changes in employment or marital status - that lead to switching schools (Welsh, 2017), while the latter is related to the student's experiences in school. To account for this, I create two outcomes of student mobility. First, I create an outcome of mobility that includes all non-structural moves, including those potentially motivated by family-related factors or school-related factors. I then create an outcome of non-structural moves that takes the value of one if the student moved for reasons other than residential mobility.[48]

I analyse the intervention impacts on both outcomes in various windows of analysis and observe that the size of the reduction in the likelihood of student mobility is similar regardless of the definition of the mobility outcome, suggesting that the LATE are mainly capturing a reduction in school-related moves. The estimated coefficients range from 1.5 and 2.4 percentage points, in most windows of analysis, which correspond, relative to the comparison group, to a 20 percent reduction in non-structural moves and a 40 percent reduction in non-structural moves that exclude residential mobility (Table 4, Panel B and C).[49]

In my main window of analysis, ± 5, I observe a 1.5 percentage point reduction only on the indicator that excludes non-residential moves, which is mainly driven by the schools located in the coastal regions of Peru. In this region, I observe that the likelihood of mobility dropped by 4 percentage points (Table F.7), while there were no effects among the schools located in the highlands and the jungle. A similar pattern is observed in other windows of analysis. This is an interesting finding as relative to the highlands and the jungle, the coastal region of Peru has a larger number of schools per square kilometre, and as a result, if the students (or parents) wanted and decided to switch schools it is more likely they would be able to do so in the coastal region relative the other regions.

The potential channel explaining the reduction in student mobility might be related to changes in school factors that contributed to improving the students' experience of school and their psychological

[48]I am able to do this as the data allows me to observe if the student moved to schools located in the same district, or located in a different district, province or region. Peru's territory is organized into regions, and these are subdivided into provinces that are composed of districts. Assuming that moves to a school located in a different province or region are a proxy for family residential moves, I create an indicator of residential mobility and then create a variable of student mobility that excludes non-structural moves related to residential mobility.

[49]The 20 percent reduction is calculated considering that the baseline indicator of student mobility (that considers all forms of non-structural moves) is equal to 6.4 percent. If we consider instead, the baseline indicator of student mobility that excludes residential mobility (that is equal to 4.2 percent), the reduction in student mobility would correspond to a 40 percent reduction in student mobility relative to the comparison group.

well-being. Two related pieces of evidence would support this. First, primary data collected from beneficiary schools shows that two-thirds of the school principals implemented for the first time positive discipline strategies after the intervention. This involves avoiding corporal and verbal punishment, building a sense of community, and using effective communication to deal with misconduct. Moreover, administrative data on school practices related to the management of violence suggests that school principals of the beneficiary schools implemented more of these practices (Table F.6). Even though this data is based on self-reports and might be noisy, it suggests that there was a shift towards school practices that signal the school's commitment to eliminating violence in schools.

Second, reporting or speaking up about experiences of violence has been found to improve psychological well-being through feelings of safety, self-protection, protection of others and redistribution of justice (Skogan 1984).[50] Students in the beneficiary were likely to have experienced these benefits considering that the intervention increased the likelihood of reporting experiences of violence in school.

Therefore, in a broader sense and drawing on the voice and exit framework described by Hirschman (1972), the reduction in student mobility can be interpreted to reflect the well-being benefits that arise from providing space to speak up about violence, even if violence levels do not change in the short term (Table 3).

[50] We have to keep in mind that well-being gains might be negatively affected by the lack of actions post-reporting. However, the reduction in student mobility would suggest that this was not the case in Peru.

Table 4: IV Estimates: Likelihood of Dropout and Mobility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Dropout								
IV - LATE	-0.0005	-0.0036	0.0024	-0.0019	-0.0017	-0.00619*	0.0003	-0.0041
	(0.00522)	(0.00539)	(0.00424)	(0.00436)	(0.00330)	(0.00342)	(0.00294)	(0.00314)
ITT	-0.000193	-0.00205	0.00152	-0.00112	-0.000894	-0.00329*	0.0000239	-0.00207
	(0.00318)	(0.00316)	(0.00264)	(0.00251)	(0.00192)	(0.00179)	(0.00160)	(0.00149)
F-stat	34.45	34.28	74.19	58.28	143.5	91.94	182.8	182.8
Anderson-Rubin Test P-values	0.924	0.674	0.800	0.870	0.808	0.183	0.469	0.469
N	403040	401412	690018	686583	978580	972834	1298759	1291918
Panel B: Mobilitty								
IV - LATE	-0.0244*	-0.0243*	-0.0145	-0.0137	-0.0162**	-0.0156*	-0.0142**	-0.0149*
	(0.01390)	(0.01350)	(0.00922)	(0.0103)	(0.00759)	(0.00827)	(0.00703)	(0.00798)
ITT	-0.0156*	-0.0151*	-0.00925	-0.00768	-0.00988**	-0.00871**	-0.00808**	-0.00745*
	(0.00860)	(0.00822)	(0.00566)	(0.00567)	(0.00454)	(0.00441)	(0.00391)	(0.00386)
F-stat	34.45	34.28	74.19	58.28	143.5	91.94	182.8	104.8
Anderson-Rubin Test P-values	0.192	0.186	0.194	0.112	0.0934	0.143	0.108	0.126
N	403040	401412	690018	686583	978580	972834	1298759	1291918
Panel C: Mobilitty (excluding residential mobility)								
IV - LATE	-0.0237*	-0.0242**	-0.0151*	-0.0147	-0.0148**	-0.0146**	-0.0137**	-0.0143**
	(0.01230)	(0.01200)	(0.00836)	(0.00930)	(0.00650)	(0.00717)	(0.00592)	(0.00682)
ITT	-0.0149*	-0.0149**	-0.00956*	-0.00835	-0.00903**	-0.00805**	-0.00775**	-0.00712**
	(0.00771)	(0.00741)	(0.00517)	(0.00520)	(0.00391)	(0.00385)	(0.00330)	(0.00329)
F-stat	34.45	34.28	74.19	58.28	143.5	91.94	182.8	104.8
Anderson-Rubin Test P-values	0.117	0.118	0.180	0.195	0.0614	0.11	0.0639	0.0886
N	403040	401412	690018	686583	978580	972834	1298759	1291918
Window of analysis	3	3	5	5	7	7	9	9
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School and Individual Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Exposure to 2018 intervention covariate	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating equation 4.3 for the windows of ± 3, ± 5, ± 7 and ± 9, including an interaction between the running variable and the treatment dummy (instrumenting for this term with an interaction term between the running variable and the eligibility dummy that takes the value of one if a school falls below the threshold). All The analysis is at the level of the students. The first row of each panel shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3 Learning

To study the short-term effects on learning I restrict my sample to secondary schools as MINEDU only administers the national standardized tests to students enrolled in second grade of secondary. Table 5 summarizes the ITT and LATE on math and language test scores. The estimated coefficients suggest that the intervention did not have any effects on learning. The estimated coefficients are positive, suggesting improvements in learning, but these are noisy and sensitive to the inclusion of lagged test scores (see columns (3) and (6).[51] It is important to highlight that in my context it might be too soon to detect any impacts on learning as the national standardized tests were administered in November, and the intervention was implemented between May and October. Unfortunately, data limitations, do not allow me to explore medium- or long-term effects in the Peruvian context.[52]

[51] I observe a drop in the coefficients despite the fact that the placebo estimates indicate no discontinuity in baseline test scores around the threshold (Figure 5 and Table F.2).

[52] In 2020 and 2021, due to the COVID-19 pandemic, MINEDU did not administer national standardized test scores. Moreover, after 2021, MINEDU administered the national standardized test only in a random sample of schools, from which only a small fraction matches my sample of analysis.

Table 5: IV Estimates: Learning

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Panel A: Math Scores						
IV - LATE	0.147	0.126	0.0138	0.192	0.172	0.0530
	(0.106)	(0.105)	(0.0587)	(0.120)	(0.119)	(0.0648)
ITT	0.0936	0.0803	0.00887	0.117	0.106	0.0298
	(0.0682)	(0.0675)	(0.0377)	(0.0751)	(0.0748)	(0.0402)
F-stat	71.56	72.91	73.22	22.57	26.51	25.64
Anderson-Rubin Test P-values	0.171	0.236	0.814	0.275	0.326	0.396
N	59678	59550	59252	59678	59550	59252
Panel B: Language Scores						
IV - LATE	0.119	0.107	-0.0250	0.143	0.135	-0.00741
	(0.112)	(0.106)	(0.0581)	(0.130)	(0.121)	(0.0639)
ITT	0.0759	0.0686	-0.0159	0.0885	0.0841	-0.00651
	(0.0718)	(0.0685)	(0.0372)	(0.0818)	(0.0769)	(0.0402)
F-stat	71.67	73.04	72.31	22.57	26.51	25.99
Anderson-Rubin Test	0.292	0.318	0.668	0.557	0.545	0.660
N	59705	59576	59278	59705	59576	59278
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	Yes	No	Yes	Yes
Lagged Test Scores	No	No	Yes	No	No	Yes
Exposure to 2018 intervention covariate	No	Yes	Yes	No	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating equation 4.3 for the window of ±5. The analysis is at the level of the students. The first row of each panel shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Robustness checks

I analyse the internal validity of the results by estimating several robustness checks.

Alternative Windows. There is a trade-off between power and specification error: in smaller windows of analysis (or bandwidths) the sample size is smaller and specification error is less likely, while in larger windows of analysis, the sample size increases but specification error is more likely. The main results presented in this paper have focused on a ±5 window of analysis. In the Appendix G, Table G.1, I present the results for the violence-related outcomes for all possible windows of analysis. Results remain similar and the first stage remains strong in alternative windows between ±4 to 8. Similarly, when analyzing the LATE on school mobility at different windows of analysis, I observe that the estimates remain statistically significant and similar in magnitude in most windows of analysis (see Appendix G, Table G.2). In the window of analysis of ±2, the results are less precise than those reported for larger windows since the sample size is reduced by at least one-quarter relative to the other windows of analysis.[53]

Functional Form. I control for local linear and quadratic polynomials[54] and add interaction terms between the running variable and the treatment dummy, instrumenting for this with interactions

[53] At all windows, the pre-treatment outcomes vary smoothly around the cutoff, except for the outcome measuring the likelihood of reporting violence that is statistically significant at the 10 percent level in the ±8 window.

[54] Following Gelman and Imbens (2019), I do not consider higher order polynomials. The authors advise not to use high-degree polynomials and give three reasons for this. First, relative to the weights based on local linear or quadratic regressions, higher-degree polynomials, in some applications, take extreme values. Second, they illustrate three applications that show that the estimated coefficients are sensitive to polynomials higher than p=2. Third, they indicate that higher-degree polynomials can produce confidence intervals that can lead to misleading inference.

between the running variable and the eligibility dummy (see Appendix G, Table G.4). The magnitude and direction of the estimated coefficients remain similar across all specifications, except in the model of quadratic polynomials with interaction terms that include covariates. In this case, I observe that the estimates lose statistical significance, and the first-stage becomes weak.

Fixed Effects: I re-estimate the specifications without including the LEMO fixed effects to allow for cross-country comparisons and observe that the results remain similar. See Table G.6.

Placebo Thresholds: I explore whether there are discontinuities in treatment at alternative thresholds and do not observe a statistically significant jump in the probability of treatment when assuming alternative cut-off points. See Appendix G, Figures G.1a and G.1b, where the grey line represents the true cut-off point, while the red line represents the placebo thresholds.

Alternative Estimation Models: The outcome variables used in this study include a non-negative count variable, as well as binary variables. A concern is that the linear model used to estimate the intervention impacts might not provide the best fit over all values of the explanatory variables. For non-negative limited dependent variables, such as the number of reports of violence, the alternative is to model the expected value of the dependent variable as an exponential function; and, for binary outcomes, the alternatives include a logistic or probit model to measure the ITT and a bivariate probit model to measure the IV estimates. I estimate the program impacts using these alternative estimation models and observe that results remain overall similar. See Appendix H.

Exposure to 2018 intervention: Few schools were exposed to a similar intervention in 2018. Considering this, I run a placebo regression in which I use as a dependent variable the treatment status in 2018 and I observe there is no jump at the discontinuity (Figure I.1, Appendix I). This confirms that treatment status in 2018 is independent to eligibility in 2019. In addition to this, I estimate my main specification in a sample that drops the schools exposed to the intervention in 2018. I analyze the LATE on the violence-related outcomes and observe that estimates remain similar in size (with some variation in precision) to the case when we use the whole sample. When analyzing the effects of the TA on student mobility, I observe that even though the coefficients remain positive, the magnitude of the coefficients shrinks and the standard errors increase (see Appendix I). This could be either because the results in school mobility were driven by the schools exposed also to the activities of the 2018 intervention, or because I lose statistical power to detect impacts after dropping 25 percent of the schools in the sample that included 46 percent of the individual level observations.

7 Conclusion

School Violence is a worldwide phenomenon affecting almost a third of students aged 13 to 15 years. Extensive empirical evidence has shown the negative effects of being a victim, a bystander, and a perpetrator of school violence. However, there is limited rigorous evidence on how to address this issue and the impacts of doing so. This paper narrows this research gap by studying the impacts of a Peruvian government-led intervention that aimed to improve the headmasters' skills in the management of school violence by training them on strategies to prevent, monitor, and deal with school violence. I exploit the eligibility rules used to select the beneficiary schools and use a fuzzy RDD to estimate the short-term impacts of the intervention on violence and educational-related indicators.

I start by exploring the intervention impacts on violence-related outcomes. Unique data on all the reports of school violence in public schools, allowed me to explore whether there was a change in the likelihood of reporting violence and the total number of reports. I found that the likelihood of reporting cases of violence increased by 15 percentage points and that the total number of reports of violence also increased among the eligible schools. Importantly, my findings indicate that physical incidents of violence were more likely to be reported, suggesting that the intervention affected decisions to report what could be considered more severe forms of violence.

Using unique administrative and primary data, I find suggestive evidence that the documented rise in reports of violence is explained by a shift in reporting behaviour rather than greater levels of school violence. Qualitative interviews showed that not all of the school community knew how and where to report and that there were fears and uncertainties related to reporting before the intervention. Moreover, administrative data showed that the treated schools worked on more school practices that had the potential to reduce reporting barriers, particularly information barriers and students' trust in the school's ability to take action against school violence. A finding that is particularly relevant for policymakers as barriers to reporting and, as a result, under-reporting of violence, limits the possibility of dealing with and reducing future events of violence.

Ensuring schools are safe and free from violence can directly impact students' experiences in school and their educational decisions. I explored this and observed that the intervention did not lead to any changes in dropout, but that it influenced decisions on whether to stay or move to a different school. I find that the likelihood of student mobility decreased by around 20 percent among treated schools, effects that would suggest improvements in the school environment and potential psychological well-being gains coming from feeling safer in school. I also explored if the intervention had any short-term effects on students learning, but found that it did not.

The results presented in this paper call for creating a research agenda to disentangle how best to prevent and manage school violence. It is important to assess the cost-effectiveness between whole-school interventions (that target both the students and the school staff), student-only interventions, and school staff-only interventions. The results of this paper suggest that a nationwide staff-only intervention can have effects similar to student-only small-scale interventions, but a careful assessment of costs is needed to understand what actions can be feasible to implement at a national level. Moreover, further research is needed to examine the link between having better school environments (free of violence) and learning outcomes, as the prevalence of school violence is a key piece of the learning crisis puzzle. Finally, this paper is only able to analyse the short-term effects of the intervention. Future research should also aim to investigate the medium- and long-term effects of policies targeting the prevention and management of violence in schools.

References

Aboagye, R. G., A.-A. Seidu, C. Adu, A. Cadri, D. O. Mireku, and B. O. Ahinkorah (2021). Interpersonal violence among in-school adolescents in sub-Saharan Africa: Assessing the prevalence and predictors from the Global School-based health survey. *SSM - Population Health 16*, 100929.

Andrews, I., J. H. Stock, and L. Sun (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics 11*(1), 727–753.

Bacher-Hicks, A., J. Goodman, J. G. Green, and M. K. Holt (2022). The COVID-19 Pandemic Disrupted Both School Bullying and Cyberbullying. *American Economic Review: Insights 4*(3), 353–370.

Baker-Henningham, H., Y. Scott, M. Bowers, and T. Francis (2019). Evaluation of a Violence-Prevention Programme with Jamaican Primary School Teachers: A Cluster Randomised Trial. *International Journal of Environmental Research and Public Health 16*(15), 2797.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy 76*(2), 169–217. Publisher: University of Chicago Press.

Bloom, N., R. Lemos, R. Sadun, and J. V. Reenen (2015). Does Management Matter in schools? *The Economic Journal 125*(584), 647–674. Number: 584.

Bradshaw, C. P., T. E. Waasdorp, and P. J. Leaf (2015). Examining variation in the impact of school-wide positive behavioral interventions and supports: Findings from a randomized controlled effectiveness trial. *Journal of Educational Psychology 107*(2), 546–557.

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

Burdick-Will, J., K. M. Nerenberg, J. A. Grigg, and F. Connolly (2021). Student Mobility and Violent Crime Exposure at Baltimore City Public Elementary Schools. *American Educational Research Journal 58*(3), 602–634.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs: Robust Nonparametric Confidence Intervals. *Econometrica 82*(6), 2295–2326.

Card, D. and L. Shore-Sheppard (2004). Using discontinuous eligibility rules to identify the effects of the federal medicaid expansions on low-income children. *The review of economics and statistics 86*(3), 752–766.

Carson, D. C., F.-A. Esbensen, and T. J. Taylor (2013). A Longitudinal Analysis of the Relationship Between School Victimization and Student Mobility. *Youth Violence and Juvenile Justice 11*(4), 275–295.

Chávez, C., V. Cebotari, M. J. Benítez, C. F. Hiu, and J. Zapata (2020). School-Related Violence in Latin America and the Caribbean: Building an Evidence Base for Stronger Schools. *Innocenti Working Paper 2021-02*, 68.

Contreras, D., G. Elacqua, M. Martinez, and Miranda (2016). Bullying, identity and school performance: Evidence from Chile. *International Journal of Educational Development 51*, 147–162.

Cortes, K. I. and B. Kochenderfer-Ladd (2014). To tell or not to tell: What influences children's decisions to report bullying to their teachers? *School Psychology Quarterly 29*(3), 336–348.

Cunha, F., Q. Hu, Y. Xia, and N. Zhao (2023). Reducing Bullying: Evidence from a Parental Involvement Program on Empathy Education. Technical Report w30827, National Bureau of Economic Research, Cambridge, MA.

Delprato, M., K. Akyeampong, and M. Dunne (2017). The impact of bullying on students' learning in Latin America: A matching approach for 15 countries. *International Journal of Educational Development 52*, 37–57.

Devries, K. M., L. Knight, J. C. Child, A. Mirembe, J. Nakuti, R. Jones, J. Sturgess, E. Allen, N. Kyegombe, J. Parkes, E. Walakira, D. Elbourne, C. Watts, and D. Naker (2015). The Good School Toolkit for reducing physical violence from school staff to primary school students: a cluster-randomised controlled trial in Uganda. *The Lancet Global Health 3*(7), e378–e386.

Doleac, J. L. and J. B. Carr (2016). The geography, incidence, and underreporting of gun violence: New evidence using ShotSpotter data. Technical report.

Dunne, M., R. Sabates, C. Bosumtwi-Sam, and A. Owusu (2013). Peer Relations, Violence and School Attendance: Analyses of Bullying in Senior High Schools in Ghana. *Journal of Development Studies 49*(2), 285–300.

Eriksen, T. L. M., H. S. Nielsen, and M. Simonsen (2014). Bullying in elementary school. *49*(4), 839–871. Number: 4.

Eslea, M. and J. Rees (2001). At what age are children most likely to be bullied at school? *Aggressive Behavior 27*(6), 419–429.

Espelage, D. L., S. Low, J. R. Polanin, and E. C. Brown (2013). The Impact of a Middle School Program to Reduce Aggression, Victimization, and Sexual Violence. *Journal of Adolescent Health 53*(2), 180–186.

Evans, D. K., S. Hares, P. A. Holland, and A. M. Acosta (2023). Adolescent Girls' Safety In and Out of School: Evidence on Physical and Sexual Violence from Across Sub-Saharan Africa. *Journal of Development Studies 59*(5), 739–757. Publisher: Taylor & Francis Journals.

Fabbri, C., K. Rodrigues, B. Leurent, E. Allen, M. Qiu, M. Zuakulu, D. Nombo, M. Kaemingk, A. D. Filippo, G. Torrats-Espinosa, E. Shayo, V. Barongo, G. Greco, W. Tol, and K. M. Devries (2021). The EmpaTeach intervention for reducing physical violence from teachers to students in Nyarugusu Refugee Camp: A cluster-randomised controlled trial. *PLOS Medicine 18*(10), e1003808. Publisher: Public Library of Science.

Fryer, R. G. (2017). Management and Student Achievement: Evidence from a Randomized Field Experiment. Technical Report Working Paper 23437.

García-Moreno, C. (2005). *WHO multi-country study on women's health and domestic violence against women: initial results on prevalence, health outcomes and women's responses*. Geneva: WHO. OCLC: 500508319.

Gasper, J., S. DeLuca, and A. Estacion (2012). Switching Schools: Revisiting the Relationship Between School Mobility and High School Dropout. *American Educational Research Journal 49*(3), 487–519.

Gelman, A. and G. Imbens (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics 37*(3), 447–456.

Gubbels, J., C. E. van der Put, and M. Assink (2019). Risk Factors for School Absenteeism and Dropout: A Meta-Analytic Review. *Journal of Youth and Adolescence 48*(9), 1637–1667.

Gutierrez, I. A., O. Molina, and H. Ñopo (2018). Stand Against Bullying: An Experimental School Intervention. Technical Report IZA DP No. 11623.

Hahn, J., P. Todd, and W. Klaauw (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica 69*(1), 201–209.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2004). Disruption versus Tiebout improvement: the costs and benefits of switching schools. *Journal of Public Economics 88*(9-10), 1721–1746.

Heckman, J. J., J. E. Humphries, S. Urzua, and G. Veramendi (2011). The effects of educational choices on labor market, health, and social outcomes. Technical Report 2011-002, Human Capital and Economic Opportunity Working Group. Issue: 2011-002.

Hepburn, L. and M. Miller (2012). Bullying and Suicidal Behaviors Among Urban High School Youth. *Journal of Adolescent Health*, 3.

Hinduja, S. and J. W. Patchin (2010). Bullying, Cyberbullying, and Suicide. *Archives of Suicide Research*, 17.

Hirschman, A. O. (1972). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.

Iyer, L., A. Mani, P. Mishra, and P. Topalova (2012). The Power of Political Voice: Women's Political Representation and Crime in India. *American Economic Journal: Applied Economics 4*(4), 165–193.

Karmaliani, R., J. McFarlane, H. M. A. Khuwaja, Y. Somani, S. S. Bhamani, T. Saeed Ali, N. Asad, E. D. Chirwa, and R. Jewkes (2020). Right To Play's intervention to reduce peer violence among children in public schools in Pakistan: a cluster-randomized controlled trial. *Global Health Action 13*(1), 1836604.

Kelly, D. R. (2017). Methods for Reducing Violence in Schools: A Systematic Review. *Journal of Educational and Developmental Psychology 7*(1).

Kim, Y. S., Y.-J. Koh, and B. Leventhal (2005). School bullying and suicidal risk in Korean middle school students. *Pediatrics 115*(2), 357–363.

Knight, L., E. Allen, A. Mirembe, J. Nakuti, S. Namy, J. C. Child, J. Sturgess, N. Kyegombe, E. J. Walakira, D. Elbourne, D. Naker, and K. M. Devries (2018). Implementation of the Good School Toolkit in Uganda: a quantitative process evaluation of a successful violence prevention program. *BMC Public Health 18*(1), 608.

Kolesár, M. and C. Rothe (2018). Inference in Regression Discontinuity Designs with a Discrete Running Variable. *American Economic Review 108*(8), 2277–2304. arXiv: 1606.04086.

Kärnä, A., M. Voeten, T. D. Little, E. Poskiparta, A. Kaljonen, and C. Salmivalli (2011). A Large-Scale Evaluation of the KiVa Antibullying Program: Grades 4-6: Evaluation of KiVa Antibullying Program. *Child Development 82*(1), 311–330.

Lalive, R., J. Van Ours, and J. Zweimüller (2006). How Changes in Financial Incentives Affect the Duration of Unemployment. *Review of Economic Studies 73*(4), 1009–1038.

Leaver, C., R. F. Lemos, and D. Dillenburg Scur (2019). Measuring and Explaining Management in Schools : New Approaches Using Public Data. Technical Report WPS9053, The World Bank. Issue: WPS9053.

Lee, D. S. and D. Card (2008). Regression discontinuity inference with specification error. *Journal of Econometrics 142*(2), 655–674.

Lee, D. S. and T. Lemieux (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature 48*(2), 281–355.

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2022). Valid t-Ratio Inference for IV. *American Economic Review 112*(10), 3260–3290.

Lee, S., C.-J. Kim, and D. H. Kim (2015). A meta-analysis of the effect of school-based anti-bullying programs. *Journal of Child Health Care 19*(2), 136–153.

Lemieux, T. and K. Milligan (2008). Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics 142*(2), 807–828.

Limber, S. P., D. Olweus, W. Wang, M. Masiello, and K. Breivik (2018). Evaluation of the Olweus Bullying Prevention Program: A large scale study of U.S. students in grades 3–11. *Journal of School Psychology 69*, 56–72.

Lleras-Muney, A. (2005). The Relationship between Education and Adult Mortality in the United States. *The Review of Economic Studies 72*, 189–221.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

McCrary, J. and H. Royer (2011). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American Economic Review 101*(1), 158–195.

Menesini, E. and C. Salmivalli (2017). Bullying in schools: the state of knowledge and effective interventions. *Psychology, Health & Medicine 22*, 240–253.

Miller, T. W. and R. F. Kraus (2008). School-Related Violence: Definition, Scope, and Prevention Goals. In T. W. Miller (Ed.), *School Violence and Primary Prevention*, pp. 15–24. New York, NY: Springer New York.

Muralidharan, K. and A. Singh (2020). Improving Public Sector Management at Scale? Experimental Evidence on School Governance India. Technical Report w28129, National Bureau of Economic Research, Cambridge, MA.

Nocentini, A. and E. Menesini (2016). KiVa Anti-Bullying Program in Italy: Evidence of Effectiveness in a Randomized Control Trial. *Prevention Science 17*(8), 1012–1023.

Oliver, C. and M. Candappa (2007). Bullying and the Politics of 'Telling'. *Oxford Review of Education 33*(1), 71–86. Publisher: Taylor & Francis, Ltd.

Olweus, D. (2005). A useful evaluation design, and effects of the Olweus Bullying Prevention Program. *Psychology, Crime & Law 11*(4), 389–402.

Oreopoulos, P. (2006). Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review 96*(1), 152–175.

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics 91*(11-12), 2213–2229.

Pescosolido, B. A. (1992). Beyond Rational Choice: The Social Dynamics of How People Seek Help. *American Journal of Sociology 97*(4), 1096–1138.

Pogarsky, G., S. P. Roche, and J. T. Pickett (2018). Offender Decision-Making in Criminology: Contributions from Behavioral Economics. *Annual Review of Criminology 1*(1), 379–400.

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

Rees, D. I., J. J. Sabia, and G. Kumpas (2022). Anti-Bullying Laws and Suicidal Behaviors Among Teenagers. *Journal of Policy Analysis and Management 41*(3), 787–823. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pam.22405.

Richardson, D. and C. Fen Hiu (2018). Developing a Global Indicator on Bullying of School-Aged Children. Innocenti Working Papers 2018/11, UNICEF. Series: Innocenti Working Papers Volume: 2018/11.

Rivers, I., V. P. Poteat, N. Noret, and N. Ashurst (2009). Observing bullying at school: The mental health implications of witness status. *School Psychology Quarterly 24*(4), 211–223.

Romero, M., J. Bedoya, M. Yanez-Pagans, M. Silveyra, and R. de Hoyos (2021). School Management, Grants, and Test Scores: Experimental Evidence from Mexico. Technical report, The World Bank.

Rumberger, R. W. and K. A. Larson (1998). Student Mobility and the Increased Risk of High School Dropout. *American Journal of Education 107*(1), 1–35.

Sarzosa, M. and S. Urzúa (2021). Bullying among adolescents: The role of skills. *Quantitative Economics 12*(3), 945–980. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE1215.

Schwartz, A. E., L. Stiefel, and S. A. Cordes (2017). Moving Matters: The Causal Effect of Moving Schools on Student Performance. *Education Finance and Policy 12*(4), 419–446. _eprint: https://direct.mit.edu/edfp/article-pdf/12/4/419/1692037/edfp_a_00198.pdf.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics 69*(1), 99.

Skogan, W. G. (1984). Reporting Crimes to the Police: The Status of World Research. *Journal of Research in Crime and Delinquency 21*(2), 113–137. Publisher: SAGE Publications Inc.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

Stock, J. and M. Yogo (2005). *Testing for Weak Instruments in Linear IV Regression*, pp. 80–108. New York: Cambridge University Press.

Strøm, I. F., S. Thoresen, T. Wentzel-Larsen, and G. Dyb (2013). Violence, bullying and academic achievement: A study of 15-year-old adolescents and their school environment. *Child Abuse & Neglect 37*(4), 243–251.

Sulak, T. N., T. F. Saxon, and D. Fearon (2014). Applying the Theory of Reasoned Action to Domestic Violence Reporting Behavior: The Role of Sex and Victimization. *Journal of Family Violence 29*(2), 165–173.

Thistlethwaite, D. L. and D. T. Campbell (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology 51*, 309–317. Place: US Publisher: American Psychological Association.

Tomczyk, S., G. Schomerus, S. Stolzenburg, H. Muehlan, and S. Schmidt (2020). Ready, Willing and Able? An Investigation of the Theory of Planned Behaviour in Help-Seeking for a Community Sample with Current Untreated Depressive Symptoms. *Prevention Science 21*(6), 749–760.

UNESCO (2019). Behind the numbers: Ending school violence and bullying. Technical report, The United Nations Educational, Scientific and Cultural Organization.

Varhama, L. M. and K. Björkqvist (2005). Relation between School Bullying during Adolescence and Subsequent Long-Term Unemployment in Adulthood in a Finnish Sample. *Psychological Reports 96*(2), 269–272.

Welsh, R. O. (2017). School Hopscotch: A Comprehensive Review of K–12 Student Mobility in the United States. *Review of Educational Research 87*(3), 475–511.

Wolke, D., W. E. Copeland, A. Angold, and E. J. Costello (2013). Impact of bullying in childhood on adult health, wealth, crime, and social outcomes. *Psychological Science 24*(10), 1958–1970.

Wolke, D. and S. T. Lereya (2015). Long-term effects of bullying. *Archives of Disease in Childhood 100*(9), 879–885.

Xie, M. and E. P. Baumer (2019). Crime Victims' Decisions to Call the Police: Past Research and New Directions. *Annual Review of Criminology 2*(1), 217–240.

Appendices

A Programs about Prevention of School Violence

Program	Countries	Program Approach	Description Intervention	Few references	Method	Findings
Olweus Bully Prevention Program (OBPP)	Several countries inluding United States and Norway	Whole School Approach	The program includes several components targeting the school staff, the students, and the community. School-level components include: establishing a Bullying Prevention Coordinating Committee (BPCC), holding staff discussion group meetings, introducing school rules against bullying, and reviewing and refining the supervisory system of schools.	Olweus (1994)	ANOVA (time-comparison treated schools)	Reduction in bullying victimization rate. Higher peer-reported assisting.
			Classroom-level components include: enforcing school-wide rules against bullying, holding regular (weekly) class meetings to discuss bullying and related topics, and holding class-level meetings with the parents. Individual-level components include: supervising students activities, ensuring that all staff intervene on the spot when bullying is observed, and meeting with students involved in bullying (and with parents of involved students).	Olweus and Limber (2010)	Literature rewview of papers studying the effect of OBPP.	Reduction in bullying victimization rate.
			Community-level components include: involving community members in the BPCC, developing community partnerships to support the program, and helping to spread anti-bullying messages of best practices in the community.	Limber et al (2008)	Multilevel regression analysis	Increase in students' expressions of empathy with bullied peers. Decrease in intentions to join in bullying.
KiVa anti-bullying	Several countries inluding Finland and Italy.	Whole School Approach	The program includes several components targeting the school staff, the students, and the parents. Teachers: the teachers receive training and support to implement the program.	Kärnä et al. (2011)	Multilevel regression analysis (random assignment of schools to treated and control arm).	Reduction in bullying victimization rate.
			Parents: the parents also receive information about bullying. Students: teachers carry out 20 hours of lessons to discuss with students what bullying is, its different forms, consequences and how individuals and groups can reduce it. The lessons also focus on fostering social skills, learning about emotions, respecting others, being part of a team and group dynamics. The lessons use role-play, video-clips about bullying, and group work and written tasks.	Nocentini and Menesini (2016)	Multilevel regression analysis (random assignment of schools to treated and control arm).	Reduction in bullying victimization rate.
				Williford et al. (2012)	Semi-estructural Model (SEM) (random assignment of schools to treated and control arm).	Reduction in bullying victimization rate. Reduction in anxiety levels.
Schoolwide Positive Behaviour Interventions & Support (SWPBIS)	United States	Student Only Approach	The program includes several components targeting the students, but teachers receive training as they are responsible for implementing the intervention. It is a noncurricular behavioural prevention strategy that follows a three-tiered process. Tier 1 includes school-wide components, while Tier 2 and Tier 3 are targeted to students at risk. The adoption of the program takes 2 to 3 years.	Bradshaw, Waasdorp, and Leaf 2015	RCT. Latent Profile Analysis.	Lower likelihood of discipline deferrals.
			Teachers: the teachers receive training and support to implement the program. Students: Tier 1 activities involve defining the student behavioural expectations, monitoring, and rewarding these expectations for all students across non-classroom and classroom settings. The reward system aims to provide a visible acknowledgement of the desirable or appropriate social behaviour.	Horner et al (2009)	Randomized, wait-list control effectiveness trial	Higher perceived safety of the school. Higher proportion of students achivieng state reading assessment standards.
				Flannery et al (2013)	Multilevel latent growth model	Lower likelihood of discipline deferrals.
Student Success Through Prevention	United States	Student Only Approach	The program includes several components targeting the students, but teachers receive training as they are responsible for implementing the intervention. Teachers: the teachers receive training to implement the program. The intervention consisted of weekly lessons (15 sessions lasting 50 minutes) aimed to foster socio-emotional skills, communication skills, and problem-solving skills to prevent bullying. The lessons combine whole-class instruction, group discussion, interactive activities, and individual work.	Espelage et al (2013)	RCT	Deecrease in the likelihood of self-reporting physical aggression.
AntibuLllying LAWS	United States	-	Anti-Bullying Laws (ABLs) include laws that included one of the following requirements: (i) provide written records of bullying and how each incident is resolved; (ii) implement strict investigatory procedures for bullying incidents; (iii) implement graduated sanctions for bullying; (iv) offer training to teachers, staff, and parents; and (v) clearly define the behaviours that constitute bullying. The laws were classified as strong or weak. Strong laws included 3 of the 5 requirements, while weak laws included less than 3.	Ress et al 2022	Differences in Differences	Decrease bullying victimization. Decrease depression and suicidal ideation.

Program	Countries	Program Approach	Description Intervention	Few references	Method	Findings
Skill Based Violence Prevention Program (VPP)	United States	Student only approach	The program activities are offered to students. The intervention has 12 sessions for students, 2 evaluation sessions and 2 planning sessions with teachers and school principals. The sessions include topics on self-concept (identifying positive traits), group dynamics, communication and conflict resolution skills, peer support against violence, and the development of classroom norms around conflict. The lessons are delivered by independent facilitators and the content of the intervention is taught using role plays, games, and project-based learning.	Thompkins et al (2014)	Hierarchical linear modeling (before and after comparison for participants and non participants).	Improvement of academic self-concept Increase in the use of conflict resolution strategies
Raising Voices Uganda - The Good School Toolkit	Uganda	Whole School Approach	The intervention included around 60 activities directed to the staff, students, and administration. It includes topics related to facilitating reflection on experiences of violence, providing knowledge on alternatives to punitive discipline, and encouraging the creation of plans, goals, and self-monitoring progress of the school goals.	Knight et al (2018) Devries et al (2015)	RCT Sample: 42 schools.	Lower likelihood of experiencing physical violence from staff.
Stand Against Bullying	Peru	Student only approach	The program activities are offered to students enrolled in urban schools. The intervention had two components. The first component focused on increasing awareness among students about the negative consequences of bullying and encouraging them to stand against this problem, and the second one focused on promoting the use of a Government online platform system to report violence.	Gutierrez et al (2018)	RCT Sample: 66 schools.	Increase in the likelihood of reporting school violence. Decrease in depression. Reduction in the likelihood of dropout and mobility. Increase in test-scores.
Parental Involvement Program in Empathy Education	China	Parents Approach	The intervention aimed to foster empathy in middle schoolers by educating and coaching their parents.	Cunha et al. (2023)	RCT	Reduced bullying. Increased parental time in empathy-building activities (i.e., watching movies or reading articles on empathy). Improved children pro-social and empathetic behaviours.
Right to Play	Pakistan	Student only approach	The intervention consisted of 103 play-based learning activities with a specific goal. After the game, the coaches hired to deliver the intervention led a three-step discussion following the formula Reflect-Connect-Apply, which involved reflection on the activity and how it made participants feel or what had been learned from it, discussion connecting this to daily life, and application more broadly to other circumstances.	Karmaliani et al. (2020)	RCT Sample: 40 schools.	Decrease bullying victimisation. Decrease corporal punishment. Decrease depression.
EmpaTeach	Tanzania	School-staff only approach	The intervention consisted on 12 self-guided sessions. The training included the following topics empathy-building exercises and on group work to learn and practice self-regulation techniques, strategies to promote wellbeing, positive disciplinary methods, and classroom management strategies.	Fabbri et al. (2021)	RCT Sample: 27 schools.	No changes in violence levels.
Violence-Prevention Programme	Jamaica	School-staff only approach	12 hours of teacher training and in-class surpport from the facilitators. The training covered the use of positive and proactive strategies to promote childrens positive behaviour and prevent negative behaviour.	Baker-Henningham et al. (2019)	RCT Sample: 14 schools.	Reduction in violence perpetrated by teachers.

Notes: there are other programmes implemented in the high-income countries such as Breaks are Better and Safe School Health Students (Madreleski et al, 2012; Sprague et al (2007). The papers describe the type of interventions and descriptive statistics about the programme. In Latin America, several government interventions have been made without a causal evaluation of the effects (Chavez et al, 2020).

B Measurement: Outcome Variables and Covariates

Variable	Definition	Source
Outcomes		
Likelihood of reporting violence	Dummy variable that takes the value of 1 if at least one event of school violence was reported, and 0 otherwise.	SISEVE data
Number of reports of violence	Sum of the reports of violence per school, including reports of any form of violence: physical, psychological, or sexual.	SISEVE data
Likelihood of student mobility	Student level: I create an indicator at the student level and at the school level. The indicator at the student level is a dummy variable that takes the value of 1 if the student is enrolled at one school during the academic year t, and enrols in a different school for the academic year t+1. I do not consider the structural moves that are required when a student needs to transition to another school because their current school does not offer the educational level they need to enrol to. In Perú this is common for transitions between primary and secondary school. Moreover, the indicator does not consider moves that occur due to school closure. Considering this, as defined by Welsh (2017), the indicator can be viewed as an indicator of non-structural mobility. That is, moves that occur when the student could have, in theory, stayed at their previous school. School level: I construct the school annual rate of mobility, which measures the proportion of students who move to a new school in the subsequent academic year.	Student Census
Likelihood of student dropout	Student level: I create an indicator at the student level and the school level. The indicator at the student level is a dummy variable that takes the value of 1 if the student enrols in the academic year t, but does not enrol in the academic year t+1, leaving the school before completing his/her studies. Taking into account that in Peru the academic year starts in March and finishes in December, a student drops out if, for example, he/she enrols in 2019 academic year, but leaves school before completing his/her studies and does not enrol in school in 2020. School level: I also construct the school annual rate of dropout, which measures the proportion of students who drop out in a single year without completing their studies.	Student Census
Math and Language Test-scores	Math and language standardized test-scores.	ECE data
Main Covariates		
Treatment Status	Dummy variable takes the value of 1 if the school was targeted to receive the intervention, and zero otherwise.	List of targeted schools (Ministry of Education)
Eligibility Dummy	Dummy variable takes the value of 1 if the school is located below the threshold rule of 9, and zero otherwise.	School Census
Running Variable	Discrete running variable normalized to zero. Each value indicates the position of the school in a ranking that is based on a weighted function of the distance of the adjacent school to the nucleo school, and the number of enrolled students.	School Census
Baseline Covariates		
School access to basic services	Dummy variable takes the value of 1 if the school has access to water, sanitation and electricity, and zero otherwise.	School Census
School infrastructure index (material of construction)	Continuous index about the material of construction of walls, roofs and floors. It is created using principal component analysis.	School Census
School Principal chosen by meritocracy	Dummy variable takes the value of 1 if the school principal was chosen meritocratically (permanent contract), and zero otherwise.	School Census
Proportion of teachers chosen by meritocracy	Proportion of teachers chosen meritocratically (permanent contract).	School Census
Proportion of parents with secondary education or more	Aggregate variable: proportion of parents that have secondary education or more. Individual level variable: dummy variable that takes the value of 1 if the parent has secondary education or more, and zero otherwise.	School Census
Secondary level	Dummy variable takes the value of 1 if the school has secondary level of education (grade7-11), and zero otherwise.	School Census
Student sex	Dummy variable takes the value of 1 if the student is a male, and zero otherwise.	School Census
Student age	Student age.	School Census
Other		
Total tasks related to the intervention	The School Census collects data on a diverse set of tasks. I create a discrete variable on the number of tasks related to the intervention. The min number of tasks is zero and the max is six.	School Census
Index of perceptions of school violence	Factor score(s) from a set of questions designed by the Ministry of education (11 statements). 5 statements ask about perceptions of violence teacher to students, while 6 statements ask about perceptions of violence between students. For each statement, the respondent indicated the frequency of exposure: (1) never, (2) one or more times annually, (3) one or more times monthly, (4) one or more times weekly. I also build separate indicators of perceptions of violence teacher to student, as well as perceptions of violence between students.	ECE data

C School Principals Survey

The school principals of the 2,650 beneficiary schools were invited to respond to an online school survey between July 1st and September 1st 2020, 7 months after the intervention.[55] In addition to the email invitation to respond to the survey, all survey participants were contacted by phone to inform them about the survey objectives and to offer the option of responding to the survey on the phone. Offering the alternative of a phone survey was particularly important in this context as it was not expected that everyone would have had access to the internet. In total, 1,235 schools responded to the survey: 54% of secondary schools with primary and secondary level and 29% of primary schools. 59 School Principals refused to answer, 83 initiated the survey but completed less than 50% of the

[55]The online survey was implemented between July 1st and September 1st of 2020, 6 months after 2019 intervention. I used Survey Solutions Software of the World Bank to implement the survey. The online platform provided information on the survey status for each participant: i) survey assigned but not initiated; ii) survey assigned and in progress; and iii) survey completed. Participants with one of the first two survey statuses received mail reminders of the survey every three days and were called up to 8 times over multiple days.

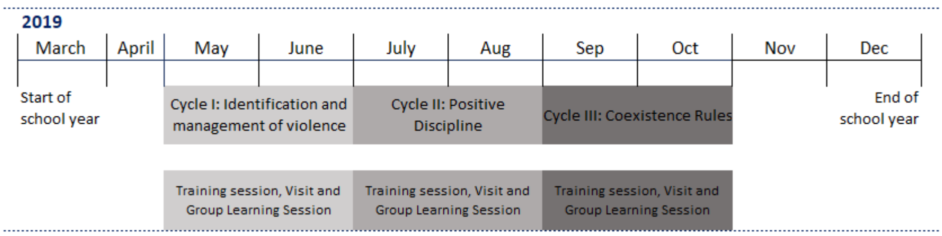
survey and the remaining did not respond to the survey, had phone numbers that were not answered, were non-existent or went to voicemail.

The School Survey data is mainly representative of urban schools that offer secondary levels of education. The survey collected data on the following variables: individual characteristics of the school principal, characteristics of the school (e.g., management index), exposure to the technical assistance, practices or tasks performed before and after the intervention, knowledge on the topics covered during the technical assistance, perceptions of the main impacts of the technical assistance, among others. This paper mainly uses information about the practices or tasks performed before and after the intervention.

D Curriculum of the Technical Assistance

The technical assistance was structured in three cycles[56] (figure D.1). Each cycle covered a new topic and included three main activities: a training session, a visit and a group learning session. The training sessions were executed at the LEMO or at an alternative venue and consisted of a detailed review of concepts, strategies and guidelines related to the management of school violence and school coexistence. The visits were executed at each school. During each visit, the LEMO Facilitator reviewed the topics of the training session and discussed the doubts or questions of the school. Finally, the group learning sessions were executed in the nucleo schools among sub-groups of 4 targeted schools (1 nucleo school and 3 adjacent schools)[57]. During these sessions, the schools shared their experiences working on the topic that was discussed during the training session.

Figure D.1: Cycles of the Technical Assistance



- *Cycle I: Identification and management of school violence.* The training topics covered: i) what is school violence and how to identify the presence of violence in the school; ii) protocols and guidelines about how to manage cases of school violence by type of violence (verbal, physical and sexual) and type of perpetrator (student and school staff); iii) platforms to register cases of violence[58]; iv) how to use the online platform SíSeVe to monitor and manage the cases of violence

[56]Section written based on the in-depth interviews and meetings executed with the division of School Management (Calidad de Gestión Escolar) at the Ministry of Education and the material used to provide the technical assistance

[57]Each LEMO was responsible for 12 schools: 3 nucleo and 9 adjacent schools. Therefore, they had 3 sub-groups of schools composed of 4 schools each.

[58]In 2018, MINEDU published protocols and guidelines for responding to school violence (Decreto Supremo 004-2018-MINEDU). Based on these protocols, the LEMO Facilitators trained the school principals on the steps to follow depending on the form of violence and perpetrator.

registered by students, confidants of the victim and bystanders. The training also highlighted the importance of informing about the reporting platforms to all the school community.

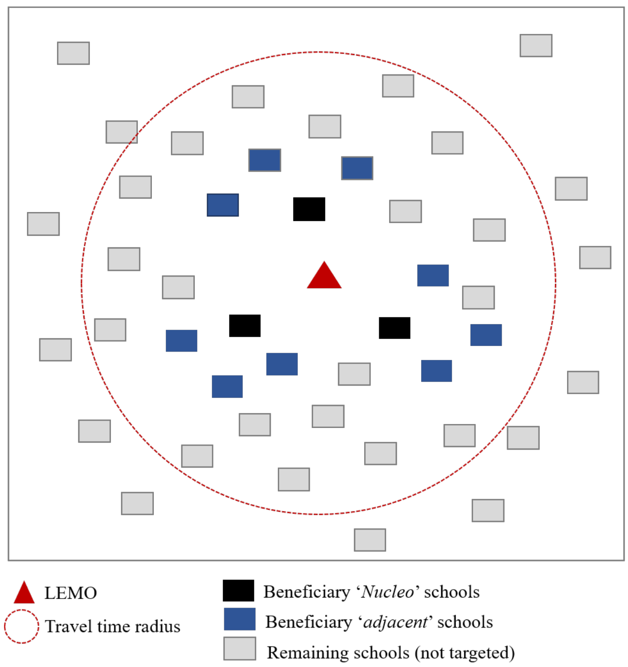
- *Second Cycle: Positive Discipline..* The training topics included: i) challenges in discipline management; ii) what is positive discipline, its principles and benefits for student development; iii) strategies to use positive discipline and establish corrective measures without using punitive discipline.
- *Third Cycle: Coexistence Rules.* The training topics included guidelines to develop coexistence rules for all the schools and for each classroom. The training highlighted that the coexistence rules had to be developed with the school community and that they had to be published in a visible location within the school and the classroom.

E Methodology to Estimate the Running Variable

To select beneficiary schools, the Ministry of Education classified schools into two groups (represented in Figure E.1):

- Nucleo schools: in each LEMO, the Ministry selected 3 schools with the highest incidence of violence, highest number of enrolled students and lowest distance to the LEMO.
- Adjacent schools: in each LEMO, the Ministry selected 9 schools. Selection of these schools depended on the schools' distance to the Nucleo schools[59] and the number of enrolled students. The Ministry started by selecting the schools that were closest in distance to one of the Nucleo schools[60], and then checked the number of enrolled students in each school to prioritize bigger schools, particularly among neighbouring schools. Therefore, in each LEMO, they selected the top 9 schools that were closer in distance to one of the Nucleo schools and that had the bigger number of enrolled students.

Figure E.1: Selection of Beneficiary Schools



Note: The figure illustrates the selection criterion, where the 3 schools with the highest incidence of violence, highest number of enrolled students and lowest distance to the LEMO (represented in the red triangle) were chosen to be the Nucleo schools (represented in the black rectangle), and, within each LEMO and for each Nucleo, the schools that were closest in distance to the Nucleo schools and that had the highest number of enrolled students were chosen to be the Adjacent schools (represented in the blue rectangle)

I will focus the analysis on the adjacent schools as the eligibility rules create an exogenous variation that allows me to estimate the impact of the TA in these schools. Another important reason to

[59]The distance variable was chosen by MINEDU as the intervention required the adjacent schools to travel to the Nucleo schools to attend the group-session component of the TA and because being located closer to the Nucleo school could minimize the likelihood of low participation rates. The second variable - number of enrolled students - was chosen with the aim of reaching more students.

[60]The selection based on distance was done by looking at a map and eyeballing which schools were closest to each Nucleo school.

focus on the adjacent schools is that 90 percent of Nucleo schools were targeted to receive another intervention at the end of the school year, making it harder to disentangle the effect of the TA for these schools. Even though MINEDU did not officially create a ranking of Adjacent schools, during in-depth interviews they explained that the selection process mimicked a ranking procedure under which the top 9 schools in each LEMO were assigned to receive the intervention. MINEDU only kept a record of the schools that were targeted to receive the TA. Therefore, to study the potential exogenous variation generated by the eligibility criteria, I estimate the ranking of schools as follows:

Part I. Distance ranking

1. Estimate the distance to each Nucleo school. I estimate distances using Vicenty (1975) formula. This method calculates the distances in kilometres between a pair of latitude and longitude points assuming an oblate sphere or an ellipsoidal model of the Earth.[61] Considering that I have three Nucleo schools per LEMO, I calculate the distances between all the schools in my sample to each of the Nucleo schools to identify to which Nucleo each school would belong based on the estimated distance.

2. Rank the schools based on their distance to each Nucleo school. I created three rankings of distance per LEMO: one per Nucleo school within each LEMO. Importantly, in 12 percent of cases[62], schools fulfil the distance eligibility criterion in more than one Nucleo school. In other words, it is possible that a school is located close to more than one Nucleo school. In these cases, I impose an excluding restriction so that if a school is eligible in one of the Nucleo schools, the school isn't considered in the ranking of distance in another Nucleo school.[63]

   The distance ranking will have values between 1 and Z, where higher values in the ranking represent the schools that are located further away from the Nucleo School.

Part II. Population ranking

3. Rank the schools based on the population or number of enrolled students. I ranked the schools based on the number of enrolled students before the intervention, where each adjacent school in the ranking will have values between 1 and Z, where higher values in the ranking represent the schools that have a lower number of enrolled students.

Part III. Distance and population ranking

4. Assign weights to the *distance* and *population of students* ranking. Qualitative interviews with the civil servants who designed the intervention revealed that the importance given to the distance and population variable varied between LEMO, mainly depending on the density and

[61] I used the Stata command called *geodist* to estimate the distances.

[62] 12 percent from a sample of schools that are within the top 30 of schools based on distance. If we consider all schools in the sample, the percentage increases to 24 percent

[63] To account for the fact that some schools can be eligible for more than one Nucleo, I start by estimating the ranking of distance in one of the Nucleo schools - e.g., Nucleo 1-, then I estimate the ranking of distance in the subsequent Nucleo - e.g., Nucleo 2- but excluding the schools that were already in the top 3 and finally I estimate the ranking of distance in the remaining Nucleo - e.g., Nucleo 3 - but excluding the schools that were already in the Top 3 in the other rankings. I also account for the fact that a school can't be part of the control group in one of the Nucleo if the school was in the top 3 or assigned to treatment in one of the other Nucleo.

dispersion of schools by LEMO. Considering this, I explore 23 different weighting schemes or weight combinations, where $W_{distance}$=[0,1] and $W_{population}$=[0,1]. This means that for each school I estimate a score for each of the 23 weighting schemes following equation E.1.[64]. See the full range of weights in Table E.1.

$$Score_{ijw} = RankDistance_{ij}W^{w}_{distance} + RankPopulation_{ij}W^{w}_{population}, \tag{E.1}$$
$$\text{where i=1 to N school, j=1 to 221 LEMO and w=1 to 23 weighting schemes}$$

5. Rank schools based on the *distance* and *population of students* weighted ranking. Using the score obtained from estimating equation E.1 for each of the 23 weighting schemes, I rank schools in ascending order, where values between 1 and 9 represent the eligible schools and 10 to Z represent potential control schools. For each LEMO, I use the weighting scheme that yields the highest predictability rate or, in other words, the scheme that predicts more closely the Ministry of Education's official selection of beneficiary schools.[65] Finally, I normalize the created ranking to zero (the running variable) and generate the eligibility dummy variable that takes the value of 1 if the value of the running variable is below zero and 0 otherwise.

I will now provide more information about the weighting schemes chosen in **step 5**. Through in-depth interviews, the Ministry of Education shared that the distance criterion was the main criterion and that the population criterion was used to prioritize between neighbouring schools. This is consistent with the fact that in 51 percent of the LEMO, my estimation procedure suggests distance weights are equal to or above 0.90 (Table E.1). For a quarter of the LEMO, my estimation procedure chooses distance weights between 0.85 and 0.70, and only in 1 percent of the LEMOs, the distance weights are equal to or below 0.10. Overall, in most of the LEMOs, distance was the predominant criterion for selecting beneficiary schools. However, the number of enrolled students also played an important role. Because of this, I do not consider alternative running variables that would either only use the distance criterion or a unique set of weights for all the LEMOs in the country.[66]

[64]It is important to note (as discussed in step 2), that for each LEMO I have 3 rankings of distance (one per Nucleo school within each LEMO). Therefore, I estimate equation E.1 for all the weighting schemes for each of the 3 distance rankings (meaning that I estimate the equation 69 times). This exercise will result in three scores per weighting scheme. I then choose the minimum score for each of the weighting schemes to be able to have a unique score per weighting scheme per LEMO. Choosing the minimum score ensures prioritizing the schools that are closer in distance to the Nucleo schools and that have a higher number of enrolled students. Moreover, this procedure ensures that each Nucleo School is allocated 3 Adjacent Schools.

$$min(ScoreNucleo^{1}_{ijw}, ScoreNucleo^{2}_{ijw}, ScoreNucleo^{3}_{ijw}),$$
$$\text{where i=1 to N school, j=1 to 221 LEMO and w=1 to 11 weighting schemes.}$$

[65]I do this by comparing the number of eligible schools predicted by my algorithm, and the number of schools selected to be treated by the Ministry.

[66]In fact, the first-stage F-statistic of a regression that only uses the distance criterion to create the running variable is below 10, signalling the weakness of the instrument.

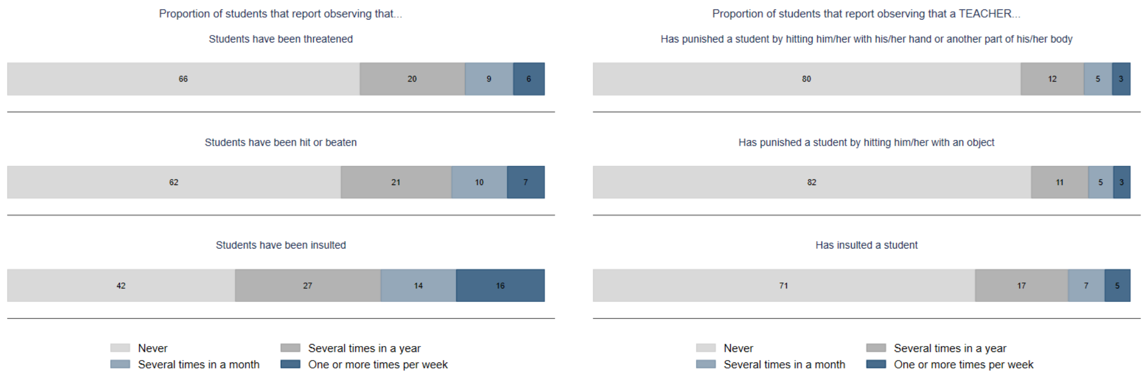
Table E.1: Number and Proportion of LEMO by Weighting Scheme

Distance Weight	Population Weight	Frequency	Percentage	Cumulative Percentage
1	0	63	28.51	28.51
0.98	0.02	5	2.26	30.77
0.95	0.05	17	7.69	38.46
0.92	0.08	14	6.33	44.8
0.9	0.1	14	6.33	51.13
0.85	0.15	19	8.6	59.73
0.8	0.2	11	4.98	64.71
0.75	0.25	11	4.98	69.68
0.7	0.3	12	5.43	75.11
0.65	0.35	16	7.24	82.35
0.6	0.4	8	3.62	85.97
0.55	0.45	3	1.36	87.33
0.5	0.5	3	1.36	88.69
0.45	0.55	4	1.81	90.5
0.4	0.6	2	0.9	91.4
0.35	0.65	2	0.9	92.31
0.3	0.7	7	3.17	95.48
0.25	0.75	3	1.36	96.83
0.2	0.8	2	0.9	97.74
0.15	0.85	2	0.9	98.64
0.1	0.9	1	0.45	99.1
0.05	0.95	1	0.45	99.55
0	1	1	0.45	100

Notes: The table provides the number and proportion of LEMOs by weighting scheme.

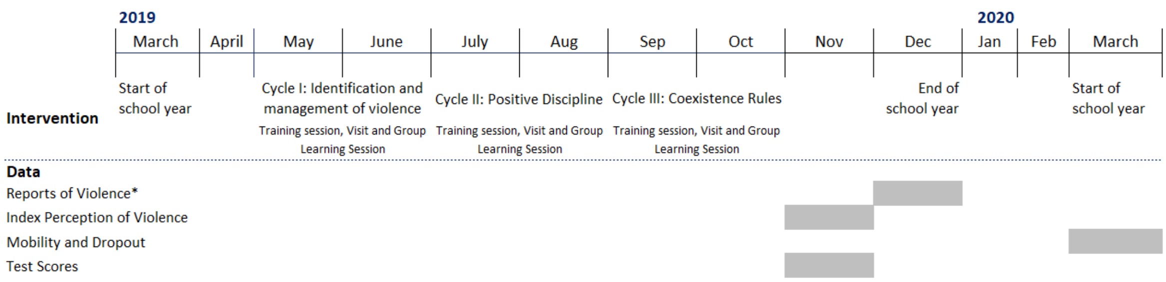
F Tables and Figures

Figure F.1: Students' perceptions of school violence



(a) Violence Between students (b) Violence teacher to student

Figure F.2: Timeline Intervention and Data Collection



Note: *Reports during all the school year, up to the end of school year.

Figure F.3: Density of Schools Around the Threshold

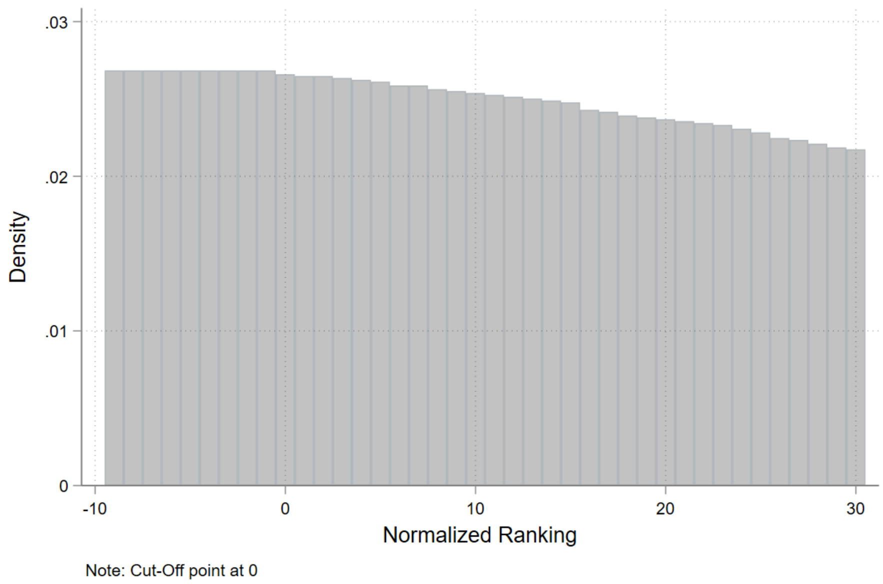


Table F.1: Normalized Differences

	Treatment Mean	Control Mean	Normalized Difference
A. Pre-treatment outcomes			
Likehood of Reporting Violence	0.233	0.166	0.166
	(0.423)	(0.373)	
Number of Reports of Violence	0.582	0.360	0.163
	(1.553)	(1.133)	
Likelihood of Student Dropout	0.019	0.02	0.005
	(0.138)	(0.140)	
Likehoof of Student Mobility	0.061	0.067	0.026
	(0.239)	(0.250)	
Math Test Scores	0.112	0.036	0.145
	(0.515)	(0.533)	
Language Test Scores	0.123	0.032	0.175
	(0.513)	(0.528)	
B. Covariates			
School has access to electricity, water and sanitation	0.648	0.6120	0.074
	(0.478)	(0.487)	
Index of School Infrastructure	0.27	0.159	0.089
	(1.234)	(1.248)	
School Principal chosen by Meritocracy	0.732	0.727	0.011
	(0.443)	(0.446)	
Proportion of teachers chosen by meritocracy	0.52	0.53	0.032
	(0.305)	(0.323)	
Proportion of parents with secondary education or more	0.616	0.641	0.104
	(0.244)	(0.242)	
School Has Secondary Level	0.577	0.418	0.322
	(0.494)	(0.493)	

Notes: The table shows the normalized differences between schools located below and above the threshold in the window of analysis of ± 5. Normalized Differences larger than 0.25 indicate that the average covariate values are different between the two groups.

Table F.2: Placebo Estimates

	N	Control Mean	Placebo
A. Pre-treatment outcomes			
Likehood of Reporting Violence	2193	0.166	0.0682
		(0.373)	(0.0588)
Number of Reports of Violence	2193	0.360	0.228
		(1.133)	(0.196)
Likelihood of Student Dropout	681774	0.020	0.00267
		(0.140)	(0.00262)
Likehoof of Student Mobility	681774	0.067	0.00320
		(0.250)	(0.00658)
Math Test Scores	59592	0.036	0.168
		(0.533)	(0.127)
Language Test Scores	59592	0.032	0.172
		(0.528)	(0.115)
B. Covariates			
School has access to electricity, water and sanitation	2193	0.612	0.0162
		(0.487)	(0.0685)
Index of School Infrastructure	2192	0.159	0.0417
		(1.25)	(0.162)
School Principal chosen by Meritocracy	2187	0.727	-0.141**
		(0.446)	(0.0693)
Proportion of teachers chosen by meritocracy	2193	0.53	-0.133***
		(0.323)	(0.0483)
Proportion of parents with secondary education or more	2193	0.64	0.01
		(0.242)	(0.0245)
School Has Secondary Level	2193	0.42	0.492***
		(0.493)	(0.0783)

Notes: The*Control Mean* shows the mean of the pre-treatment variables for schools located above the cut-off. The column *Placebo* presents the estimated coefficients and standard errors obtained after estimating equation 4.3 but using as dependent variable the outcomes and relevant covariates determined prior to the intervention. The estimates correspond to a window of analysis of ± 5. Controls excluded from the table include quadratic distance to cutoff. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table F.3: Effects on reports of violence by grade

	Primary Level	Secondary Level	All	Primary Level	Secondary Level	All
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV: Likelihood of Reporting Violence	0.204**	0.179	0.134*	0.209***	0.229*	0.146**
	(0.0792)	(0.116)	(0.0684)	(0.0799)	(0.125)	(0.0696)
F-stat	101.1	71.17	182.2	141.0	36.81	101.3
Anderson-Rubin Test	0.0141	0.173	0.0670	0.0200	0.253	0.0751
IV: Number of Reports of Violence	0.389*	1.158**	0.684**	0.395*	0.960*	0.668**
	(0.217)	(0.537)	(0.289)	(0.216)	(0.501)	(0.277)
F-stat	100.8	70.37	181.8	141.0	37.10	102.3
Anderson-Rubin Test	0.0902	0.0533	0.0250	0.180	0.149	0.0776
N	1728	1088	2186	1728	1088	2186
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Exposure to 2018 intervention covariate	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients obtained after estimating equation 4.3 for the window of ±5, controlling for covariates. Columns (4) to (6) include an interaction between the running variable and the treatment dummy (instrumenting for this term with an interaction term between the running variable and the eligibility dummy that takes the value of one if a school falls below the threshold). Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table F.4: Descriptive statistics on reports' characteristics

	Type of violence			Who reports	
	Physical	Psychological	Sexual	School Staff	Other
Treated group (below threshold)	52.27%	33.60%	15.66%	51.16%	48.84%
Comparison group (above threshold)	42.62%	41.71%	14.13%	52.46%	47.54%
p-value	0.000	0.002	0.414	0.626	0.488

Table F.5: Treatment effects by type of violence

	Likelihood of reporting	Number of Reports	Likelihood of reporting	Number of Reports
	(1)	*(2)*	*(3)*	*(4)*
Physical Violence	0.0926	0.373**	0.102	0.380**
	(0.0641)	(0.164)	(0.0653)	(0.160)
F-stat	182.2	181.8	101.3	102.3
Anderson-Rubin Test	0.172	0.0282	0.186	0.0573
Sexual Violence	0.0739	0.134	0.0690	0.123
	(0.0467)	(0.0915)	(0.0542)	(0.0885)
F-stat	205.2	181.8	101.3	102.3
Anderson-Rubin Test	0.0290	0.167	0.367	0.289
Psychological Violence	0.0592	0.227	0.0664	0.210
	(0.0509)	(0.154)	(0.0506)	(0.147)
F-stat	182.2	181.8	101.3	102.3
Anderson-Rubin Test	0.278	0.171	0.280	0.350
N	2186	2186	2186	2186
LEMO fixed effects	Yes	Yes	Yes	Yes
School Covariates	Yes	Yes	Yes	Yes
Exposure to 2018 intervention covariate	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	Yes	Yes

Notes: Columns (1) and (3) use as an outcome a dummy variable that takes the value of 1 if the school had a report of violence of a specific type, and zero otherwise. Columns (2) and (4) use as an outcome the number of reports of violence by type. The table presents the estimated coefficients obtained after estimating equation 4.3 for the window of ±5 (controlling for covariates). Columns (3) and (4) include an interaction between the running variable and the treatment dummy (instrumenting for this term with an interaction term between the running variable and the eligibility dummy that takes the value of one if a school falls below the threshold). Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table F.6: Number of school practices related to the management of school violence

	Only Primary Level	Only Secondary Level	Primary & Secondary	All	Only Primary Level	Only Secondary Level	Primary & Secondary	All
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
IV-LATE	0.219	0.994*	0.517	0.360*	0.151	0.761	0.449	0.344
	(0.386)	(0.547)	(0.350)	(0.208)	(0.393)	(0.735)	(0.363)	(0.212)
F-stat	49.97	20.63	30.92	181.9	68.97	4.092	21.54	101.6
Anderson-Rubin Test P-values	0.617	0.119	30.92	0.102	0.661	0.220	0.395	0.185
N	1098	458	630	2186	1098	458	630	2186
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exposure to 2018 intervention covariate	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	No	No	Yes	Yes	Yes	Yes

Notes: The table presents the estimated coefficients obtained after estimating equation 4.3 for the window of ±5, controlling for covariates. Columns (5) to (8) include an interaction between the running variable and the treatment dummy (instrumenting for this term with an interaction term between the running variable and the eligibility dummy that takes the value of one if a school falls below the threshold). Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table F.7: IV Estimates: Likelihood of Mobility by Region

	Coast		Jungle		Mountains	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV: Likehood of Mobility	-0.0345*	-0.0348	0.0301	0.0224	-0.00888	-0.0126*
	(0.0189)	(0.0212)	(0.0214)	(0.0190)	(0.00700)	(0.00765)
F-stat	36.24	23.00	21.48	19.91	54.02	10.94
Anderson-Rubin Test	0.0517	0.0724	0.174	0.393	0.231	0.107
IV: Likelihood of Mobility (exclusing residential mobility)	-0.0408**	-0.0410**	0.0286	0.0225	-0.00446	-0.00649
	(0.0183)	(0.0200)	(0.0191)	(0.0147)	(0.00560)	(0.00580)
F-stat	36.24	23.00	21.48	19.91	54.02	10.94
Anderson-Rubin Test	0.0140	0.0477	0.129	0.276	0.447	0.367
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School & Individual Level Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Exposure to TA 2018 Covariate	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	Yes	No	Yes	No	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating oequation 4.3 for the windows of analysis ±5. Panel A shows the coefficient estimates for the outcome of mobility that includes all forms of non-structural moves, while Panel B shows the estimated coefficients for the outcome of mobility that excludes residential moves. The analysis is at the level of the students. Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

G Robustness Checks

Table G.1: IV Estimates in different Windows: Violence Related Outcomes

Window of Analysis	N	Likelihood of Reporting Violence		Number of Reports of Violence	
		(1)	(2)	(1)	(2)
2	879	0.0259	0.0259	0.649	0.648
		(0.154)	(0.155)	(0.647)	(0.646)
3	1319	0.0338	0.0341	0.503	0.504
		(0.107)	(0.110)	(0.442)	(0.443)
4	1755	0.176**	0.175**	0.835***	0.848***
		(0.0760)	(0.0759)	(0.304)	(0.309)
5	2192	0.134*	0.146**	0.684**	0.668**
		(0.0684)	(0.0696)	(0.289)	(0.277)
6	2629	0.125*	0.134**	0.769**	0.762**
		(0.0659)	(0.0664)	(0.307)	(0.303)
7	3063	0.151**	0.153**	0.558**	0.556**
		(0.0609)	(0.0611)	(0.248)	(0.247)
8	3496	0.139**	0.140**	0.607**	0.606**
		(0.0587)	(0.0586)	(0.242)	(0.242)
9	3927	0.0784	0.0786	0.262	0.263
		(0.0536)	(0.0538)	(0.231)	(0.231)
School Covariates		Yes	Yes	Yes	Yes
Exposure to 2018 intervention covariate		Yes	Yes	Yes	Yes
LEMO fixed effects		Yes	Yes	Yes	Yes
Polynomial		p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment		No	Yes	No	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating equation 4.3. Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table G.2: IV Estimates in different Windows: Likelihood of Dropout and Mobility

Window of Analysis	N	***Likelihood of dropout***				***Likelihood of Mobility***				***Likelihood of Mobility (excluding residential mobility)***			
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
2	265114	0.0144	0.0113	0.0144	0.0114	0.0278	-0.0242	-0.0284	-0.025	-0.0269	-0.0247	-0.0274	-0.0253
		(0.0116)	(0.0115)	(0.0115)	(0.0115)	(0.0232)	(0.0209)	(0.0246)	(0.0223)	(0.0210)	(0.0190)	(0.0220)	(0.0200)
3	401412	-0.000565	-0.00378	-0.000489	-0.00364	-0.0240*	-0.0236*	-0.0244*	-0.0243*	-0.0234**	-0.0238**	-0.0237*	-0.0242**
		(0.0053)	(0.0054)	(0.0052)	(0.0054)	(0.0134)	(0.0129)	(0.0139)	(0.0135)	(0.0119)	(0.0115)	(0.0123)	(0.0120)
4	543176	0.00237	0.0001	0.0020	-0.000325	-0.0112	-0.00971	-0.0106	-0.00887	-0.0121	-0.0114	-0.0117	-0.0108
		(0.0038)	(0.0039)	(0.0043)	(0.0043)	(0.0084)	(0.0084)	(0.0094)	(0.0095)	(0.0077)	(0.0076)	(0.0080)	(0.0080)
5	686583	0.00212	-0.00217	0.00239	-0.0019	-0.0126	-0.0102	-0.0145	-0.0137	-0.0137*	-0.0123	-0.0151*	-0.0147
		(0.0042)	(0.0044)	(0.0042)	(0.0044)	(0.0084)	(0.0090)	(0.0092)	(0.0103)	(0.0077)	(0.0083)	(0.0084)	(0.0093)
6	841769	-0.000385	-0.00509	-0.00036	-0.00509	-0.0167**	-0.0155*	-0.0164**	-0.0154*	-0.0151**	-0.0141**	-0.0149**	-0.0141*
		(0.0031)	(0.0032)	(0.0031)	(0.0032)	(0.0075)	(0.0080)	(0.0076)	(0.0084)	(0.0066)	(0.0071)	(0.0065)	(0.0073)
7	972834	-0.0018	-0.00619*	-0.00173	-0.00619*	-0.0160**	-0.0150*	-0.0162**	-0.0156*	-0.0147**	-0.0141**	-0.0148**	-0.0146**
		(0.0034)	(0.0034)	(0.0033)	(0.0034)	(0.0074)	(0.0079)	(0.0076)	(0.0083)	(0.0063)	(0.0068)	(0.0065)	(0.0072)
8	1122895	-0.000111	-0.0044	-0.000238	-0.00476	-0.0131*	-0.0123*	-0.0136**	-0.0133*	-0.0120**	-0.0114*	-0.0125**	-0.0122*
		(0.0029)	(0.0029)	(0.0029)	(0.0030)	(0.0067)	(0.0071)	(0.0069)	(0.0074)	(0.0057)	(0.0061)	(0.0059)	(0.0064)
9	1291918	0.000841	-0.00339	0.000303	-0.0041	-0.0130*	-0.0132*	-0.0142**	-0.0149*	-0.0126**	-0.0128**	-0.0137**	-0.0143**
		(0.0028)	(0.0029)	(0.0029)	(0.0031)	(0.0067)	(0.0075)	(0.0070)	(0.0080)	(0.0056)	(0.0063)	(0.0059)	(0.0068)
School Covariates		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Exposure to TA 2018 Covariate		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
LEMO fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Polynomial		p=1	p=1	p=1	p=1	p=1	p=1	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment		No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating equation 4.3. Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table G.3: Violence Outcomes using a Quadratic Specification

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Likelihood of reporting violence						
IV - LATE	0.224***	0.134*	0.135**	0.261***	0.199	0.210*
	(0.0638)	(0.0686)	(0.0683)	(0.0968)	(0.124)	(0.123)
ITT	0.130***	0.0685*	0.0692*	0.118	0.0578	0.0546
	(0.0403)	(0.0374)	(0.0372)	(0.0845)	(0.0809)	(0.0790)
F-stat	206.1	183.3	183.2	1.408	1.525	1.610
Anderson-Rubin Test P-values	0.00138	0.0680	0.0641	0.00523	0.122	0.0784
N	2193	2186	2186	2193	2186	2186
Panel B: Number of reports of violence						
IV - LATE	0.939***	0.677**	0.682**	0.604	0.659	0.688
	(0.269)	(0.288)	(0.287)	(0.398)	(0.468)	(0.465)
ITT	0.546***	0.346**	0.349**	0.430	0.378	0.366
	(0.166)	(0.155)	(0.154)	(0.356)	(0.301)	(0.301)
F-stat	206.1	182.8	182.8	1.408	1.471	1.572
Anderson-Rubin Test P-values	0.00115	0.0265	0.0248	0.0121	0.169	0.149
N	2193	2186	2186	2193	2186	2186
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=2	p=2	p=2	p=2	p=2	p=2
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors for the window of ±5. Robust standard erros are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table G.4: Student Dropout and Mobility using a Quadratic Specification

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dropout						
IV - LATE	0.00224	-0.00195	-0.00210	-0.0332	-0.0306	-0.0329
	(0.00420)	(0.00439)	(0.00433)	(0.257)	(0.187)	(0.204)
ITT	0.00142	-0.00114	-0.00122	0.00333	0.00312	0.00306
	(0.00266)	(0.00257)	(0.00253)	(0.00511)	(0.00511)	(0.00511)
F-stat	120.5	109.8	109.7	0.00635	0.00907	0.00864
Anderson-Rubin Test P-values	0.596	0.658	0.630	0.381	0.548	0.479
N	690018	686583	686583	690018	686583	686583
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School and Individual Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=2	p=2	p=2	p=2	p=2	p=2
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes
Panel B: Mobilitty						
IV - LATE	-0.0146	-0.0129	-0.0132	0.0327	0.0286	0.0261
	(0.00891)	(0.00963)	(0.00967)	(0.354)	(0.266)	(0.260)
ITT	-0.00922	-0.00749	-0.00771	-0.0149	-0.0151	-0.0152
	(0.00561)	(0.00561)	(0.00562)	(0.0115)	(0.0109)	(0.0108)
F-stat	120.5	109.8	109.7	0.00635	0.00907	0.00864
Anderson-Rubin Test P-values	0.102	0.183	0.172	0.441	0.557	0.540
N	690018	686583	686583	690018	686583	686583
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School and Individual Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=2	p=2	p=2	p=2	p=2	p=2
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes
Panel C: Mobilitty (excluding residential mobility)						
IV - LATE	-0.0151*	-0.0141	-0.0144	0.0174	0.0135	0.0116
	(0.00821)	(0.00893)	(0.00895)	(0.246)	(0.180)	(0.175)
ITT	-0.00956*	-0.00822	-0.00839	-0.0142	-0.0144	-0.0145
	(0.00513)	(0.00515)	(0.00516)	(0.0103)	(0.00979)	(0.00975)
F-stat	120.5	109.8	109.7	0.00635	0.00907	0.00864
Anderson-Rubin Test P-values	0.0637	0.112	0.105	0.295	0.419	0.394
N	690018	686583	686583	690018	686583	686583
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School and Individual Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=2	p=2	p=2	p=2	p=2	p=2
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors for the window of ±5. Robust standard erros are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table G.5: IV Estimates clustering standard errors at the LEMO level or by the running variable

	(1)	*(2)*	*(3)*	*(4)*
Likelihood of Reporting Violence	0.134*	0.134***	0.146**	0.146**
	(0.0684)	(0.0499)	(0.0696)	(0.0595)
Number of Reports of Violence	0.684**	0.684***	0.668**	0.668***
	(0.289)	(0.112)	(0.277)	(0.0958)
Likelihood of Dropout	-0.00217	-0.00217	-0.00190	-0.00190
	(0.00441)	(0.00328)	(0.00436)	(0.00357)
Likehood of Mobility	-0.0102	-0.0102	-0.0137	-0.0137***
	(0.00900)	(0.00879)	(0.0103)	(0.00510)
Likelihood of Mobility (exclusing residential mobility)	-0.0123	-0.0123*	-0.0147	-0.0147***
	(0.00827)	(0.00638)	(0.00930)	(0.00425)
Math Scores	0.0138	0.0111	0.0530	0.0533
	(0.0587)	(0.0462)	(0.0648)	(0.0370)
Language Scores	-0.0250	-0.0213	-0.00741	-0.00129
	(0.0581)	(0.0411)	(0.0639)	(0.0370)
N	2186	2186	2186	2186
School Covariates	Yes	Yes	Yes	Yes
LEMO fixed effects	Yes	Yes	Yes	Yes
Exposure to 2018 intervention covariate	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	Yes	Yes
SE Clustered by	LEMO	Running Variable	LEMO	Running Variable

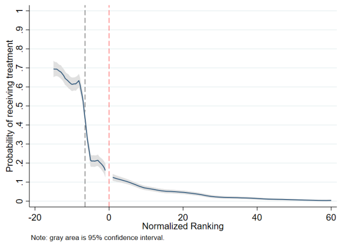
Notes: The table presents the estimated coefficients and standard errors for the window of ±5. Columns (1) and (3) show the estimated coefficients when standard errors are clustered at the LEMO level, while columns (2) and (4) show the estimated coefficients when standard errors are clustered by the running variable. *$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table G.6: IV Estimates with and without LEMO Fixed Effects

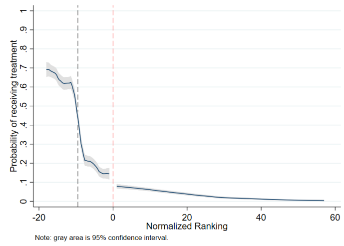
	(1)	*(2)*	*(3)*	*(4)*
Panel A: Violence Outcomes				
Likelihood of Reporting Violence	0.119*	0.134*	0.129*	0.146**
	(0.0655)	(0.0684)	(0.0664)	(0.0696)
F-stat	217.4	182.2	110.8	101.3
Anderson-Rubin Test P-values	0.0720	0.0670	0.0771	0.0751
Number of Reports of Violence	0.595**	0.684**	0.582**	0.668**
	(0.273)	(0.289)	(0.263)	(0.277)
F-stat	216.6	181.8	111.7	102.3
Anderson-Rubin Test P-values	0.0298	0.0250	0.0909	0.0776
N	2186	2186	2186	2186
Panel B: Dropout and Mobility				
Likelihood of dropout	-0.00435	-0.00217	-0.00360	-0.00190
	(0.00463)	(0.00441)	(0.00459)	(0.00436)
F-stat	100.0	114.4	73.42	58.28
Anderson-Rubin Test P-values	0.347	0.624	0.582	0.870
Likelihood of mobility (excluding residential mobility)	-0.0137	-0.0123	-0.0160*	-0.0147
	(0.00853)	(0.00827)	(0.00952)	(0.00930)
F-stat	100.0	114.4	73.42	58.28
Anderson-Rubin Test P-values	0.111	0.136	0.231	0.195
Panel C: Learning				
Math Scores	-0.0204	0.0138	0.0272	0.0530
	(0.0670)	(0.0587)	(0.0665)	(0.0648)
F-stat	76.27	73.22	42.61	25.64
Anderson-Rubin Test P-values	0.762	0.814	0.319	0.396
Language Scores	-0.0229	-0.0250	-0.000581	-0.00741
	(0.0553)	(0.0581)	(0.0547)	(0.0639)
F-stat	77.16	72.31	43.65	25.99
Anderson-Rubin Test P-values	0.682	0.668	0.607	0.660
LEMO fixed effects	No	Yes	No	Yes
School Covariates	Yes	Yes	Yes	Yes
Exposure to 2018 intervention covariate	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	Yes	Yes

Notes: Columns (1) and (3) present the estimated coefficients obtained after estimating equation 4.3 for the window of ±5 but without including LEMO fixed effects, while columns (2) and (4) include LEMO fixed effects. Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure G.1: Alternative Placebo Thresholds



(a) Threshold at 15



(b) Threshold at 18

H Alternative Estimation Models

The outcome variables used in this study include a non-negative count variable, as well binary variables. A concern is that the linear model used to estimate the intervention impacts might not provide the best fit over all values of the explanatory variables. For non-negative limited dependent variables, such as Number of Reports of Violence, the alternative is to model the expected value of the dependent variable as an exponential function; and, for binary outcomes the alternative involves using a logistic model and a bivariate probit.

- Non-negative limited dependent variables The distribution of the variable Numbers of Reports of Violence has a right-skewed distribution, that takes very few values and has a mean and median at 0. Considering this, I estimate equation H.1 using a Poisson regression, where T_{ij} is an eligibility dummy that takes the value of 1 for those schools which position in the ranking is equal or lower to the cut-off, and 0 otherwise; and, $g(ranking_j)$ corresponds to a parametric function of the running variable. Equation H.1 will gives the reduced-form intent-to-treat estimates.

$$Y_{ij} = exp(\beta + \lambda T_{ij} + g(ranking_j)) + \mu_{ij} \tag{H.1}$$

  Table H.1 presents the intent-to-treat estimates. The first row shows the results estimated using a linear model, while the second row shows the results from using a Poisson Regression. Results are not directly comparable with OLS. Yet, they indicate that the reported cases of violence were higher among the beneficiary schools by more than 100%.

- Binary dependent variables

  Table H.2 and H.3, Panel A, present the intent-to-treat estimates. The first row shows the estimates from a linear model, while the second row shows the marginal effects from a logistic model. Moreover, in Panel B, the first row shows the LATE based on a linear model, while the second row shows the estimates from a non-linear model (using a bivariate probit). We observe similar effects on the likelihood of reporting violence under the different estimation models.

Table H.1: Reports of School Violence

	(1)	(2)	(3)	(4)
ITT - OLS: Number of Reports of Violence	0.547***	0.350**	0.546***	0.349**
	(0.166)	(0.155)	(0.166)	(0.154)
ITT - POISSON: Number of Reports of Violence	0.803***	0.626***	0.797***	0.621***
	(0.209)	(0.207)	(0.207)	(0.213)
N	2193	2186	2193	2186
LEMO fixed effects	Yes	Yes	Yes	Yes
School Covariates	No	Yes	No	Yes
Exposure to TA 2018 Covariate	No	Yes	No	No
Polynomial	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	Yes	Yes

Notes: The table presents the estimated coefficients for the window of ±5. The first row shows the intent-to-treat estimates from a linear model, while the second row shows the estimates from a Poisson model. Standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regarding the likelihood of mobility we observe similar ITT estimates. However, when using a bivariate probit, I observe that estimates are sensitive to the inclusion of covariates.

Table H.2: Likelihood of reporting violence

Panel A: ITT - OLS and Logistic Model	*(1)*	*(2)*	*(3)*	*(4)*
OLS: Likelihood of reporting violence	0.130***	0.0685*	0.130***	0.0690*
	(0.0403)	(0.0372)	(0.0403)	(0.0372)
LOGIT: Likelihood of reporting violence	0.162***	0.101**	0.170***	0.109**
	(0.0472)	(0.0423)	(0.0484)	(0.0435)
Panel B: LATE - IV and BiProbit Model	*(1)*	*(2)*	*(3)*	*(4)*
IV: Likelihood of reporting violence	0.223***	0.134*	0.232***	0.146**
	(0.0638)	(0.0684)	(0.0647)	(0.0696)
BiProbit: Likelihood of reporting violence	0.2986***	0.2117***	0.2607**	0.1621**
	(0.0702)	(0.0557)	(0.1030)	(0.0775)
N				
LEMO fixed effects	Yes	Yes	Yes	Yes
School Covariates	No	Yes	No	Yes
Exposure to 2018 intervention covariate	No	Yes	No	No
Polynomial	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	Yes	Yes

Notes: The table presents the estimated coefficients for the window of ±5. Panel A The first row shows the ITT estimates from a linear model and a logistic model. Panel B shows the LATE from a linear and a non-linear model. Standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I Exposure to 2018 Intervention

Context: In 2018, MINEDU offered an intervention similar to the 2019 intervention. The intervention in 2018 was offered in 181 of the 221 LEMO to 2,605 schools in the country. In 2019, the intervention was offered in all the LEMOs to 2,655 schools. 30% of the schools that received the intervention in 2019, were also exposed to at least one activity of the 2018 intervention. Over the two years, the selection criteria considered similar variables, but the eligibility rules and the process of selection changed over time:

- In 2018, MINEDU selected between 12 to 17 schools in each LEMO. The schools were selected based on the number of enrolled students and the number of cases of violence reported in the SíSeVE platform, prioritizing the schools that were either bigger in size and/or had more reports

Table H.3: Likelihood of Student Mobility

Panel A: ITT - OLS and Logistic Model	*(1)*	*(2)*	*(3)*	*(4)*
Likelihood of school mobility				
OLS	-0.00795	-0.00599	-0.00925	-0.00768
	(0.00526)	(0.00527)	(0.00566)	(0.00567)
LOGIT	-0.00815	-0.00563	-0.00911*	-0.00688
	(0.00512)	(0.00500)	(0.00529)	(0.00514)
Likelihood of school mobility (excluding residential mobility)				
OLS	-0.00865*	-0.00719	-0.00956*	-0.00835
	(0.00478)	(0.00481)	(0.00517)	(0.00520)
LOGIT	-0.00859*	-0.00655	-0.00910**	-0.00722
	(0.00449)	(0.00440)	(0.00460)	(0.00446)
Panel B: LATE - IV and BiProbit Model				
Likelihood of school mobility				
IV	-0.0126	-0.0102	-0.0145	-0.0137
	(0.00837)	(0.00900)	(0.00922)	(0.0103)
BiProbit	-0.013***	-0.001	0.004	0.005***
	(0.00288)	(0.00184)	(0.00288)	(0.00196)
Likelihood of school mobility (excluding residential mobility)				
IV	-0.0137*	-0.0123	-0.0151*	-0.0147
	(0.00765)	(0.00827)	(0.00836)	(0.00930)
BiProbit	-0.012***	-0.006***	0.002	0.001
	(0.00217)	(0.00146)	(0.00217)	(0.00152)
LEMO fixed effects	Yes	Yes	Yes	Yes
School Covariates	No	Yes	No	Yes
Exposure to 2018 intervention covariate	No	Yes	No	Yes
Polynomial	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	Yes	Yes

Notes: The table presents the estimated coefficients for the window of ±5. Panel A The first row shows the ITT estimates from a linear model and a logistic model. Panel B shows the LATE from a linear and a non-linear model. Standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of violence. Among the schools that fulfilled one or both conditions, MINEDU prioritized the schools located near the LEMO due to both logistic and budget constraints.

- In 2019, the changes in the eligibility criteria were motivated by the changes in the activities of the intervention, particularly the inclusion of group learning sessions; and, to reach a wider range of schools, including those that based on the available data had not experienced incidents of violence. MINEDU categorized schools into two groups: Nucleo schools and Adjacent schools. Nucleo schools, similar to 2018, were selected based on the number of enrolled students, the prevalence of violence and their distance to the LEMO. Adjacent schools were selected based on the distance to the Nucleo schools and the number of enrolled students. See Appendix E for details.

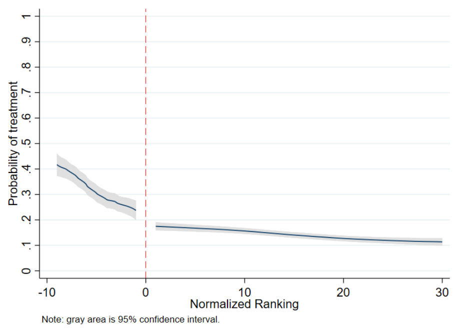
I treat the 2018 and 2019 interventions independently and focus on analysing the impact of the 2019 intervention.[67] Four reasons motivate this. First, the likelihood of treatment in 2018 is orthogonal or independent of 2019 eligibility criteria (see section 4). Second, the differences in the eligibility criteria and intervention design. Third, the eligibility criteria used to select the adjacent schools in 2019 provide an opportunity to evaluate the impact of the intervention in a causal way. Fourth, the

[67] Alternatively, I could have treated the interventions as if they were the same intervention but implemented following a staggered roll-out. The latter option would involve using a two-way-fixed effects regressions model (accounting for the new advances in the method, particularly De Chaisemartin and D'Haultfceuille (2019, 2020) suggestions on how to deal with programs where treatment switches on/off across time). However, the differences in the eligibility criteria would make the comparison challenging. Moreover, the differences in the intervention curriculum across time would also affect the interpretation of the treatment effects.

TA in 2019 was more homogeneous across the LEMO relative to the 2018 intervention. In 2018, due to logistic constraints, the intervention was not implemented in 40 LEMOs. Moreover, even though the intervention was planned to be implemented between June and November[68], in 21 LEMO the intervention started in October[69]. These LEMOs had less time to implement all programme activities and hence, as qualitative interviews revealed, in some beneficiary schools it was not possible to execute all the activities of the intervention. In 2019, the intervention was implemented between May and October, and administrative data shows that two-thirds of the LEMO delivered all program components and that the remaining LEMO implemented, on average, 7 out of the 9 program components.

Robustness checks: I run a placebo regression in which I use as a dependent variable the treatment status in 2018 and I observe there is no jump at the discontinuity (figure I.1, appendix I). This confirms that treatment status in 2018 is independent to eligibility in 2019. Moreover, as a robustness check, I drop from the sample all the schools that were exposed to 2018 TA and observe that the estimated coefficients for the violence-related outcomes remain similar. I only observe an increase in the standard errors that can be explained by the fact that the sample of schools shrinks by 25 percent. This finding would suggest that I am able to identify the LATE of 2019 intervention (see tables I.2 and I.3). When analyzing the effects of the TA on student mobility, I observe that even though the coefficients remain positive, the magnitude of the coefficients shrinks and the standard errors increase (see table I.1). This could be either because the results in school mobility were driven by the schools exposed also to the activities of the 2018 intervention, or because I loose statistical power to detect impacts after dropping 25 percent of the schools in the sample that included 46 percent of the individual level observations.

Figure I.1: Probability of being Treated in 2018



[68]The academic year in Perú starts in March and finishes by mid-December.

[69]Based on administrative data, in 21 LEMO the first visit to the schools was executed in October

Table I.1: Likelihood of reporting violence

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Sample						
IV - LATE	0.223***	0.133*	0.134*	0.232***	0.144**	0.146**
	(0.0638)	(0.0687)	(0.0684)	(0.0647)	(0.0699)	(0.0696)
ITT	0.130***	0.0678*	0.0685*	0.130***	0.0683*	0.0690*
	(0.0403)	(0.0374)	(0.0372)	(0.0403)	(0.0374)	(0.0372)
F-stat	205.2	182.3	182.2	86.18	100.5	101.3
Anderson-Rubin Test P-values	0.00145	0.0710	0.0670	0.00118	0.0800	0.0751
N	2193	2186	2186	2193	2186	2186
Panel B: Drop schools exposed to 2018 TA						
IV - LATE	0.221***	0.146*	-	0.226***	0.156*	-
	(0.0740)	(0.0832)	-	(0.0749)	(0.0833)	-
ITT	0.120***	0.0684	-	0.117***	0.0659	-
	(0.0440)	(0.0423)	-	(0.0444)	0.043	-
F-stat	110.8	90.76	-	77.06	105.3	-
Anderson-Rubin Test	0.00708	0.107	-	0.00236	0.0504	-
N	1636	1631	-	1636	1631	-
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3). Panel A includes all schools located in the window of ±5, while Panel B excludes the schools that were exposed to activities activities of 2018 intervention. The first row of each Panel shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.3: Mobility

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Sample						
IV - LATE	-0.0137*	-0.0121	-0.0123	-0.0151*	-0.0144	-0.0147
	(0.00765)	(0.00827)	(0.00827)	(0.00836)	(0.00926)	(0.00930)
ITT	-0.00865*	-0.00711	-0.00719	-0.00956*	-0.00820	-0.00835
	(0.00478)	(0.00482)	(0.00481)	(0.00517)	(0.00519)	(0.00520)
F-stat	117.8	114.7	114.4	74.19	60.97	58.28
Anderson-Rubin Test P-values	0.0713	0.141	0.136	0.180	0.211	0.195
N	690018	686583	686583	690018	686583	686583
Panel B: Drop schools exposed to 2018 TA						
IV - LATE	-0.00824	-0.00488	-	-0.00990	-0.00723	-
	(0.0114)	(0.0119)	-	(0.0121)	(0.0126)	-
ITT	-0.00503	-0.00265	-	-0.00579	-0.00350	-
	(0.00694)	(0.00651)	-	(0.00724)	(0.00676)	-
F-stat	51.87	44.62	-	41.03	49.15	-
Anderson-Rubin Test	0.469	0.685	-	0.546	0.461	-
N	371926	369793	-	371926	369793	-
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3) in the window of ±5. Panel A includes all schools located in the window of ±5, while Panel B excludes the schools that were exposed to activities activities of 2018 intervention. In each panel, the first row shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I.2: Number of Reports of School Violence

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Panel A: All Sample						
IV - LATE	0.940***	0.679**	0.684**	0.930***	0.663**	0.668**
	(0.270)	(0.290)	(0.289)	(0.259)	(0.278)	(0.277)
ITT	0.547***	0.347**	0.350**	0.546***	0.346**	0.349**
	(0.166)	(0.155)	(0.155)	(0.166)	(0.155)	(0.154)
F-stat	205.2	181.8	181.8	86.18	101.3	102.3
Anderson-Rubin Test P-values	0.00117	0.0267	0.0250	0.00257	0.0810	0.0776
N	2193	2186	2186	2193	2186	2186
Panel B: Drop schools exposed to 2018 TA						
IV - LATE	0.812***	0.652**	-	0.805***	0.645**	-
	(0.282)	(0.305)	-	(0.273)	(0.294)	-
ITT	0.441***	0.305**	-	0.443***	0.306*	-
	(0.165)	(0.153)	-	(0.168)	(0.155)	-
F-stat	110.8	90.53	-	77.06	106.3	-
Anderson-Rubin Test	0.00798	0.0477	-	0.0140	0.102	-
N	1636	1631	-	1636	1631	-
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	Yes	No	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	No	No	Yes
Polynomial	p=1	p=1	p=1	p=1	p=1	p=1
Interaction: Running Var*Treatment	No	No	No	Yes	Yes	Yes

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3). Panel A includes all schools located in the window of ±5, while Panel B excludes the schools that were exposed to activities activities of 2018 intervention. The first row of each Panel shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$