Improving School Management of Violence: Evidence from a Nation-wide Policy in Peru.

Gabriela Smarrelli[1]

[1]Department of International Development, University of Oxford

November, 2021

Abstract

Exposure to school violence has 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 intervention that aimed to improve the school heads' 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 37 percentage points and that the number of reports of violence also rose among eligible schools. Using unique administrative, qualitative, and primary data, I find suggestive evidence that the documented rise in reports of violence is primarily due to shifts in reporting behaviour rather than a greater incidence of school violence. Upon exploring the short-term impacts on education-related outcomes, I find the intervention reduced the student likelihood of switching schools by three percentage points. These findings add to our understanding of the benefits of investing in school staff skills that contribute to the creation of safer learning environments.

Keywords: economics of education, school management of violence, school mobility, school dropout, test-scores.

JEL Classification: I20, I29, H75.

1 Introduction

Almost a third of students aged 13 to 15 years worldwide have been victims of the harmful effects of school violence (UNESCO, 2019). This is a complex behavioural phenomenon that takes different forms - including physical, sexual, and psychological violence - and that can emerge as a set of isolated events against different victims or as repeated attacks against the same victim (the latter is what we commonly refer to as *bullying*).

School violence directly affects children's right to inclusive and equitable education, and is detrimental for human capital formation. Extensive evidence, mainly from the 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 of victimization with student dropout, student mobility, and absenteeism (Brown and Taylor 2008; Dunne et al. 2013; Carson et al. 2013; Burdick-Will et al. 2021). Brown and Taylor (2008) even show that exposure to school violence outweighs the negative effects of class size in educational attainment. Evidence also documents that experiencing school violence has lasting negative effects over the life cycle both in terms of 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 et al, 2012). Importantly, these adverse effects 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, has led to a rise in policies and programs directly targeting school violence. Since the early 2000's the United States has implemented state-specific anti-bullying laws (Ress et al, 2020) 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, there is limited rigorous evidence about the impact of these interventions, particularly beyond high-income countries. Two empirical challenges that explain this include the absence of valid comparison groups to estimate robust treatment effects and the limited availability of reported violence data and victimization surveys.

This paper aims to contribute to this knowledge gap by analysing the case of Peru, a middle-income country with one of the highest rates of victimization at school in Latin America. In 2019, the Ministry of Education (MINEDU) designed a large-scale Technical Assistance (TA) aiming to improve the management of school violence. 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 cases of violence, the adoption of response protocols, and the implementation of positive discipline strategies. In each of the provinces of the country,

[1] The following are just a few of the countries that have implemented school coexistence and anti-bullying programs: United States, Canada, Peru, United Kingdom, Finland, Denmark, Norway, Uganda (Kelly, 2017; UKAID, 2018; WHO, 2020). See appendix 8.1 for details.

MINEDU offered the TA to 12 schools: 3 nucleo schools[2] and 9 adjacent schools. In this paper, I limit the analysis to the adjacent schools mainly for two reasons. First and foremost, 90% 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 parameters relative to other alternatives.

To select the adjacent schools, MINEDU chose, in each province of the country, the top-9 schools that were closer 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 then 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 (see figure 5). I produce intention to treat (ITT) estimates and, using the eligibility dummy as an instrument for treatment, I produce instrumental variable (IV) estimates that provide the local average treatment effect (LATE) 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.*

Reporting violence. The intervention increased the likelihood of reporting by 37 percentage points. Among the schools that reported violence, some schools had already registered incidents of violence at least once before the intervention, while others registered events of violence for the first time in 2019. I explore the likelihood of reporting for both groups of schools and observe a higher likelihood of reporting among first-time reporters.

I also explore changes in the number of reports of violence and find that they increased, on average, by 1 report among eligible schools. This result is non-trivial considering that the median number of reports, among the schools that reported events of violence, was 2 reports.

One plausible explanation for these results is that they are reflecting a change in the 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). Taken together, the following pieces of evidence suggest that the intervention could have contributed to reducing the barriers to reporting.

First, the likelihood of reporting is mainly driven by the schools that might have faced the biggest barriers before the intervention. That is, the schools that reported at least one case of violence for the first time in 2019. Second, administrative and primary data show that eligible schools worked more on tasks that have the potential to contribute to reducing fears and uncertainties related to reporting. These include the dissemination of channels for reporting, the creation of spaces to discuss the topic

[2] In Spanish *'nucleo'* means centre or core.

of school violence, and the development of school coexistence rules with the school community.

An alternative story could be that the increase in reporting is reflecting an increase in violence levels. This story is perhaps less likely and difficult to study considering the issue of underreporting. 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). For instance, in Peru, administrative data, collected by MINEDU from students aged 13 to 14 years, revealed that in more than 90% of the schools that documented a case of violence for the first time in 2019, students had witnessed events of physical and psychological violence in the past. This evidence indicates that that not every event of violence is being reported and points to the importance of reducing the barriers to reporting.

Importantly, using this dataset, 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 differences in the index of perception of school violence between the schools eligible to receive the TA and those that were not. This serves as suggestive evidence that there was not an increase in violence levels due to the intervention.

Therefore, in the short term, my estimates are more likely to inform about shifts in reporting behaviour rather than informing about changes in levels of violence. This finding is particularly relevant for policymakers as underreporting limits the possibility of dealing with and reducing future events of violence[3].

Educational outcomes. Next, I turn to the education-related outcomes, and using student-level data, I study the impact of the program on student dropout, student mobility, and test scores. I observe that the intervention reduced the likelihood of student mobility by 3 percentage points. The indicator of student mobility only considers non-structural moves that occur when the student could, in theory, have stayed at their previous school. Among non-structural moves, a common reason for school mobility is residential mobility. That is, cases in which the family might move to a new province (e.g., due to divorce, 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 students' experience of school.

Regarding the other education-related outcomes, I find that the TA did not have an impact on student dropout. Moreover, the estimated coefficients on language and math test scores are positive, suggesting improvements in test scores, however these are not statistically significant. Test-scores are measured one month after the intervention so it may be too soon to measure impacts on these variables. An additional point to consider is that I might lack statistical power to measure impacts on test-scores with precision as this information is only available for secondary schools (45% of my sample).

The results are robust to changes in the functional form, the estimates remain similar at different windows of analysis and I do not observe any jumps in any of the pre-treatment outcome variables.

[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 are driven by greater reporting rather than actual increases in crime.

Therefore, it is unlikely there are serious threats to the internal validity of the estimates. It is, however, important to mention that 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 to analyse with precision whether the impacts differ by pre-treatment characteristics.

This paper contributes to the scarce literature studying the impacts of interventions targeting the phenomenon of school violence (Kelly, 2017, Chávez et al, 2021). The limited rigorous available evidence comes mainly from high-income countries that have studied the impact of state-antibullying laws (Rees et al. 2020), 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(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-middle income countries face different constraints related to the quality of the systems of education, the budget, and the culture of punitive discipline that prevails in many settings. To my knowledge, other than this paper, only two interventions targeting primary and secondary schools 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 likelihood of physical violence from staff to students, as well as the likelihood of absenteeism. In Peru, Gutierrez et al. (2018) implemented a randomized control trial in 66 urban schools in 2015 and studied the effects of a student-only intervention that consisted of providing information to the students about the negative consequences of bullying, the importance of standing against bullying, and the available reporting platforms. The authors found their intervention increased the willingness to report cases of violence and reduced the likelihood of school mobility by 2 percentage points.

This paper adds new evidence by analysing the impacts of a nationwide 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. Moreover, even though whole-school interventions are considered ideal (Lee et al. 2015), governments 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 contributes to this discussion by showing that interventions focused solely on strengthening the school heads' violence management skills can have similar effects to student-only interventions.

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 of switching schools. With regards to learning, the evidence is mixed. Similar to Devries et al. (2015) that studied the effects of a whole-

[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.

school type of intervention, I do not find evidence that the TA had short-term effects on learning. Instead, Gutierrez et al (2018) evidenced improvements in math and language test scores in the medium term. Therefore, this paper hopes to motivate further research on the short, medium, and long-term relationship between school anti-violence strategies and learning outcomes, as this is an important piece of the learning crisis puzzle (Pritchett 2013; Angrist et al. 2021).

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) show 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 to 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 34,211 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:

Census Schools Characteristics: 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 type of contract or position, by gender, educational background) and 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

[5]For 98% of the schools, I have data over the 6 years. For the remaining 2%, 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 close 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.

mobility). I also use the Student Census data of 2020, to identify whether the student enrolled to school and whether he/she move to a new school in the academic year of 2020.

SíSeVE Reports: report level data that allows to identify the number of violence reports per school, the form of violence, as well as, the age and sex of the victim and the perpetrator. The dataset also allows to observe who registered the incident of violence as cases can be reported either by the victim or a witness.

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 included a set of questions about the 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 of 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 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 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 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. This situation is less

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

[7] The ECE is only administered in schools with more than five students enrolled are not considered for the test. Before 2019, the ECE was also administered to children enrolled in primary education. Between 2006 and 2016, the ECE was also applied to students enrolled in second grade of primary - children aged 8 years old on average, and in 2016 and 2018, the ECE applied to students in fourth grade of primary, aged 10 years of old.

common: between 2014 and 2019, 3% 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 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 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 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 8.2 for details).

2.2 Primary Data

I complement the administrative data with in-depth interviews executed to MINEDU officials and to facilitators located at the Local Educational Offices (LEMO), and with unique primary survey data that was collected 4 to 6 months after the intervention. I executed two online surveys: a survey to the LEMO facilitators, in charge of implementing the TA in 2019, and a survey to school principals from the beneficiary schools[10]. The LEMO survey was responded by 80% of LEMO Facilitators, while the school principal survey was responded by 54% of secondary schools and 29% of primary schools[11]. These surveys allow me to complement the analysis with detailed descriptive statistics about the implementation process of the TA and the school practices that changed after 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 35,000 primary and secondary public schools[12]. 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

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

[9]Even though administrative data does not allow to observe the address of the student, it allows 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. I do this as residential mobility is generally associated with family related factors such as carer access to a new job or divorce.

[10]Budget constraints did not allow to extend the survey to schools in the comparison group.

[11]Overall, the LEMO Facilitators that responded the survey are similar, in average, to those that did not respond the survey. In the case of the survey to school principals, the respondents were similar, in terms of school characteristics, to those that did not respond. However, those that responded were more likely to be located in urban areas. Therefore, results are particularly representative for urban schools. See Appendix 8.3 for details.

[12]Primary levels cover 6 years of education from age 6 to 11, while secondary levels, cover 5 years from age 12 to 16. Moreover, 87% offer primary education, either in the form of multi-grade teaching (students in a classroom belong to different grades) or single-grade teaching (students in a classroom belong to the same grade), and 27% offer secondary education, all following single-grade teaching

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[13] – called SíSeVE[14].

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 (TA) that provided training to school principals and a teacher representative on the following topics: utilization of the online system to report and identify cases of violence, protocols to manage cases of violence, 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 intervention, 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 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 Hui, 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 what we commonly refer to 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)[15] and the 2013 Third Regional Comparative and Explanatory Study (TERCE)[16] show that 47% 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 online platform created by MINEDU. In addition to SíSeVe, the victims of school violence can report in person at their nearest LEMO or, since 2019, by phone[17].

The data based on SíseVE reports has allowed, for the first time in Peru, to detect events of school

[13]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)

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

[15]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.

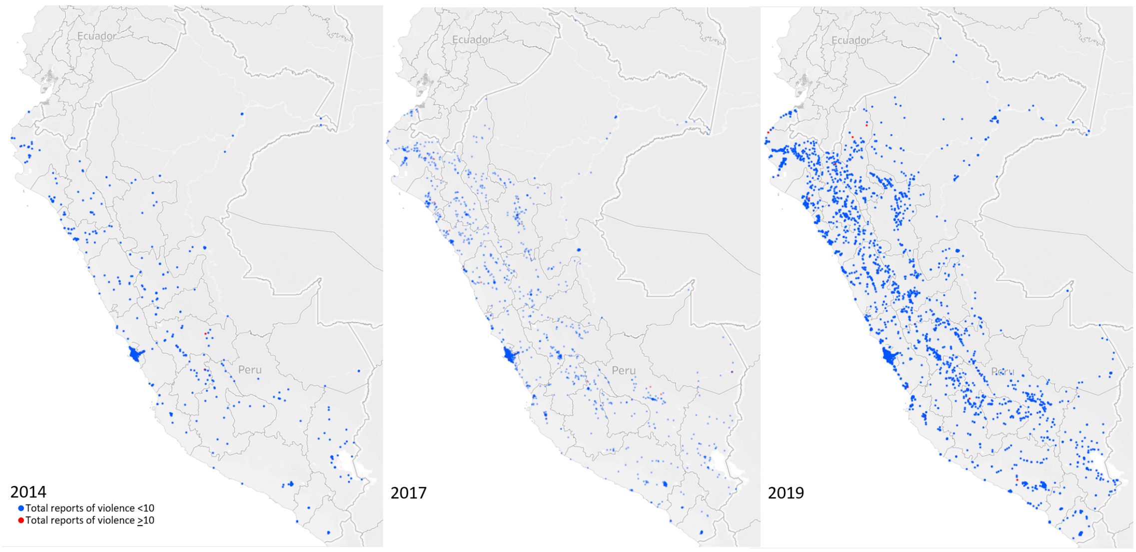
[16]Data on 4403 students aged 11 to 12 years of age. The indicator is constructed using a series of questions in which students indicate whether they have been victims of different forms of violence in the last month and whether they fear other students in the school.

[17]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.

violence[18]. In the last 6 years, half of the reports were related to incidents of physical violence, follow up by reports of psychological (30%) and sexual violence (20%). 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. School violence against girls and boys was not statistically significantly different, except in 2019 when a higher number of reports of violence against girls was registered. Overall, the most frequent form of violence against girls was psychological violence (40% of reports) follow up by one third of cases of physical violence and one third of sexual violence (includes rape, sexual assault and sexual harassment). Among boys, sexual violence is the least common form of violence, with around 5% of reports, while two thirds of the reports refer to physical violence and one third to psychological violence.

Between 2014 and 2019, 20% 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 88 reports of violence per 100 thousand enrolled students, while in 2019, the number doubled: 202 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% registered cases for the first time that year.

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



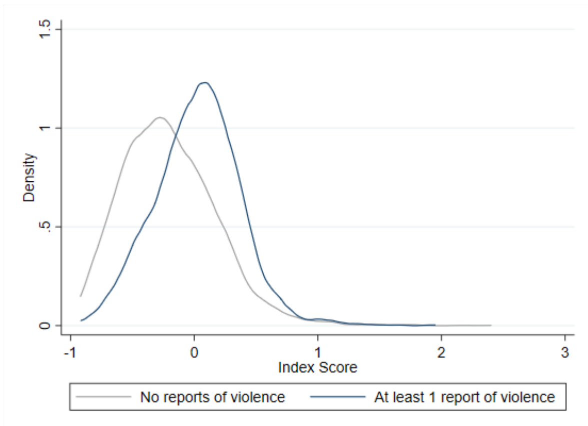
Second, **student survey data suggests that not all events of violence are reported**. Among the schools that registered cases of violence, the number of recorded reports may underestimate the true number of incidents of violence. Among the schools that did not register any case, it is uncertain whether incidents of violence occurred. Survey data collected by MINEDU in 2018 and 2019 from

[18]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% of the cases of violence in the last 6 years, were reported by family members and 50% by members of the school staff. Among the limitations, is the fact that around 50% of public primary schools and 25% of public secondary schools[19] in the country do not have access to internet connection. In these cases, reporting by phone, in-person, or using public internet booths remains an option.

students aged between 13 and 14 years old allowed me to explore this further[20]. The survey asked students[21] if they had observed or witnessed incidents of violence perpetrated by other students or the 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[22]. In 2019, 50% of surveyed students said they witnessed at least 1 event of physical and psychological violence between students, and 22% witnessed events of both physical and psychological violence from teachers to students[23]. 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 Leaving or Switching Schools

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 Perú.

[20]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

[21]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 8.6, figure 2 summarizes a few of the statements that were included in the survey

[22]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.

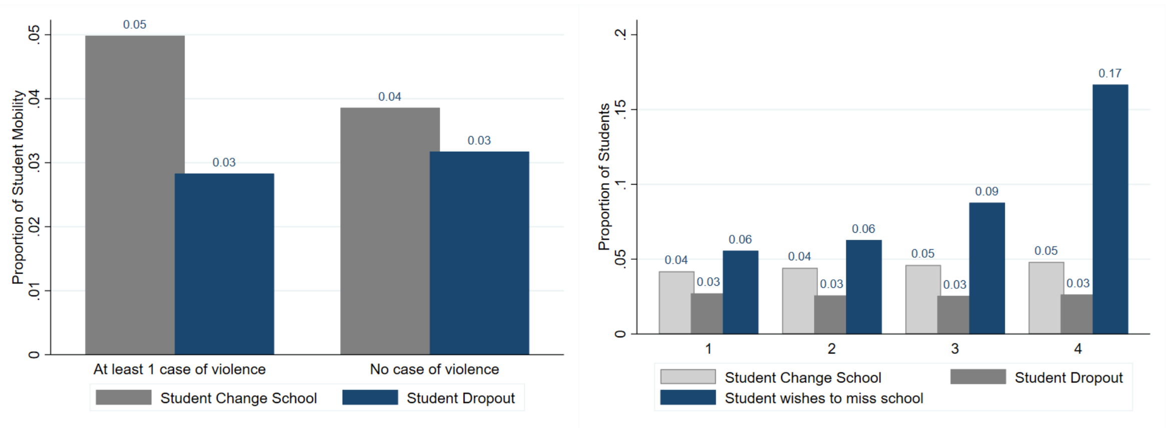
[23]This statistic is estimated by creating a dummy variable that takes the value of 1 if the student said he/she witnessed at least 1 of the statements used in the survey to identify the presence of physical and psychological violence

In the last decade, the rate of student dropout in public schools has remained stable, around 3% to 4%. Regarding student mobility, the rate of students that switched or moved[24] to new schools was between 5% and 6%. 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% and 42% 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 and perceptions of violence among secondary public schools, I observe that the rate of student mobility is slightly higher in schools that registered at least one case of violence[25]. 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[26]. 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 TA both on reporting behaviour and educational indicators.

Figure 3: School Violence and Education Indicators



(a) School reported violence

(b) Perceived presence of school violence

3.3 Technical Assistance

MINEDU designed the Technical Assistance with the aim of improving the prevention and management of school violence[27]. In each LEMO, a civil servant (from now onwards, LEMO Facilitator)

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

[25]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.

[26]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'.

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

was responsible for implementing the intervention. The TA was structured in three cycles involving three training sessions, three visits, and three group learning sessions (see appendix 8.4). Each training session lasted around 4 hours and introduced a new topic in the following order: identification, registration of cases of violence and use of the SíSeVE platform; guidelines and protocols to manage incidents 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 TA constitute 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[28] 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[29] 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.

After selecting the potential beneficiary schools, MINEDU shared the list of schools with the LEMO. The local offices had some discretionary power to provide suggestions to target other schools. However, any suggestion had to be backed up with evidence (e.g., school closure; alternative schools that also fulfilled the distance criteria but was more vulnerable) and any potential changes had to be approved by MINEDU, reducing concerns with respect to preferences towards selecting certain schools. Moreover, qualitative interviews with the LEMO confirm that the schools were not aware of the eligibility criteria and could not self-select to receive the Technical Assistance.

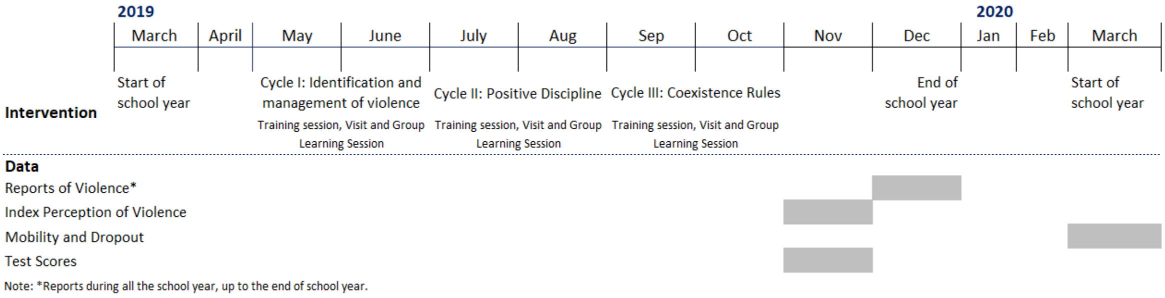
In the following section, I discuss in detail the empirical strategy that I follow to measure the

[28]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.

[29]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.

impact of the 2019 intervention. Considering that TA was implemented between May and October of 2019 and the outcomes were measured 1 to 4 months after the completion of the intervention (figure 4), 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*).

Figure 4: Timeline Intervention and Data Collection



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 TA and that is unlikely to differ from the beneficiary schools in terms of their observable and unobservable characteristics.

Eligibility rules: In each LEMO, MINEDU 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 exogeneous 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% of nucleo school 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.[30] 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[31] and that had the highest number of enrolled students. Even though MINEDU did not officially elaborate a ranking, they explained that the selection process mimicked a ranking procedure under which the top 9 schools were assigned to receive the intervention.

The ranking procedure and the top-9 threshold rule[32] provides an opportunity to analyse the impacts of the TA using a Regression Discontinuity Design (RDD). This method, introduced by Campbell and Thistlethwaite (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

[30]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.
[31]3 Adjacent Schools per Nucleo School.
[32]The decision to target 9 schools was based on budget constraints.

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 exogeneous variation generated by the eligibility criteria, I estimate the ranking of schools as follows (see Appendix 8.5 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[33].

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 11 different weighting schemes ranging between $W_{distance}$=1 and $W_{population}$=0 and $W_{distance}$=0.60 and $W_{population}$=0.40[34]. This means that for each school I create a score that is associated to each of the 11 weighting schemes following equation 4.1.[35]

$$\begin{gathered} Score_{ijw} = RankDistance_{ij} W^{w}_{distance} + RankPopulation_{ij} W^{w}_{population}, \\ \text{where i=1 to N school, j=1 to 221 LEMO and w=1 to 11 weighting schemes} \end{gathered} \tag{4.1}$$

4. *Create the ranking based on the score obtained in the previous step*: I then rank schools on 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[36]. Finally, I normalize the chosen ranking to zero and this becomes my assignment or running variable.

Figure 5 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, indicates that in 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.

[33] Considering there are 3 Nucleo Schools per LEMO, I have a total of 3 distance rankings per LEMO.

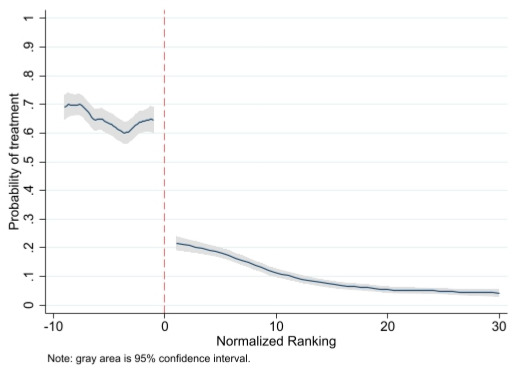
[34] I do not use weights lower than 0.4 because i) when constructing weights using regression analysis the results indicate that the average weight of the distance variable was 0.9 and ii) when repeating the latter exercise for each LEMO separately, weights below 0.4 for the distance variable were unlikely.

[35] Considering that for each school I have 3 rankings of distance, equation 4.1 is estimated for all the weighting schemes for each of the 3 distance rankings (33 times). After doing these, for each weighting scheme I chose the combination of distance ranking and population that gives the min score. This procedure ensures that each Nucleo School is allocated 3 Adjacent Schools.

[36] Predictability rate is defined as the proportion of schools assign to treatment based both on MINEDU official treatment dummy and the predicted eligibility dummy.

As a result, in 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.

Figure 5: 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 [37]. $g(ranking_j)$ corresponds to a parametric function of the running variable and λ_j represents LEMO fixed effects.

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

In the second stage, I estimate the following specification[38]:

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

where y_{ij} represents the outcome variable of interest for school i located in the province where the

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

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

LEMO j operates. $\hat{D_{ij}}$ represents the predicted probability of the treatment status dummy. Various smooth forms of $g(ranking_j)$ are considered, including linear and quadratic[39]. 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 TA 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[40].

The running variable used in this paper is discrete. This is common in other RDD applications that use, to mention a few, age (Lemieux and Milligan, 2008; Lalive, 2008), date of birth (McCrary and Royer,2003; Card and Shore-Sheppard, 2004; Oreopoulos,2006) and number of employees (Hahn, Todd and VanderKlaauw, 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 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.

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, Kolesar and Rothe (2017) find that this approach has poor coverage properties and is unable to resolve model misspecification concerns in discrete settings. The authors find that clustering by the running variable provides inappropriate narrow confidence intervals and suggest using more conservative heteroskedasticity-robust standard errors[41]. In appendix 8.9, 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 ±4, a neighbourhood that contains 1764 schools. I chose this window since it represents a middle point, without being too close or too far from the cutoff, and considering that it ensures that at least 1 school per nucleo will be included in the sample of analysis. Using a data-driven approach developed by Calonico et al. (2014)[42], 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).

[39] I also explore models that include interaction terms between the eligibility dummy and the running variable (see section 6.

[40] The addition of fixed effects allows to concentrate in within LEMO variation. I also estimate the same specification but removing the fixed effects (see section 6.

[41] Kolesar and Rothe (2017) also propose using *honest confidence intervals*. This method is recommended particularly when the number of support points close to the threshold is so small that it is not feasible to have a small bandwidth. In my setting, this could be important when using small windows of analysis. Therefore, I will also explore this in the coming months.

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

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 enrollment was registered by the schools prior to the intervention.[43,44]

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. To explore this, I plot the unconditional expectation functions of predetermined outcomes and covariates against the running variable on either side of the threshold (see figure 6 and 7). I also explore this formally by estimating equation 4.3 but using as a dependent variable the predetermined outcomes and covariates (Lee and Lemieux, 2010) (see appendix 8.6, table 2). I do not find discontinuities in any of the predetermined outcomes and in the majority of the baseline covariates. The two covariates for which I find a discontinuity are *proportion of teachers that have been chosen by meritocracy* and *secondary level*. The latter one is a dummy variable that takes the value of 1 if the school offers secondary level education. Schools in Peru can offer both primary and secondary level of education, or they might offer only primary or secondary education. The discontinuity in this variable indicates that below the threshold the proportion of schools with secondary level is higher. The fact that these two covariates are "locally" unbalanced across different windows of analysis might be a source of concern. However, even though these baseline covariates jump 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 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. (2020)[45] 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.

[43] In the few cases were the LEMO suggested modifications to the ranking, the LEMO gave specific reasons and the final decision was taken by MINEDU.

[44] 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 each 221 LEMO, there are only 221 schools located in the position 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 3 in Appendix 8.6).

[45] Lee et al (2020) 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 outcome variables

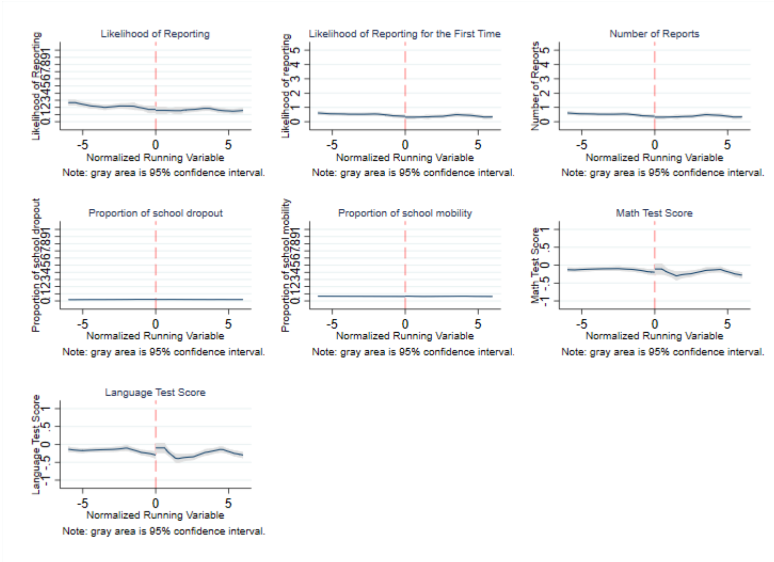
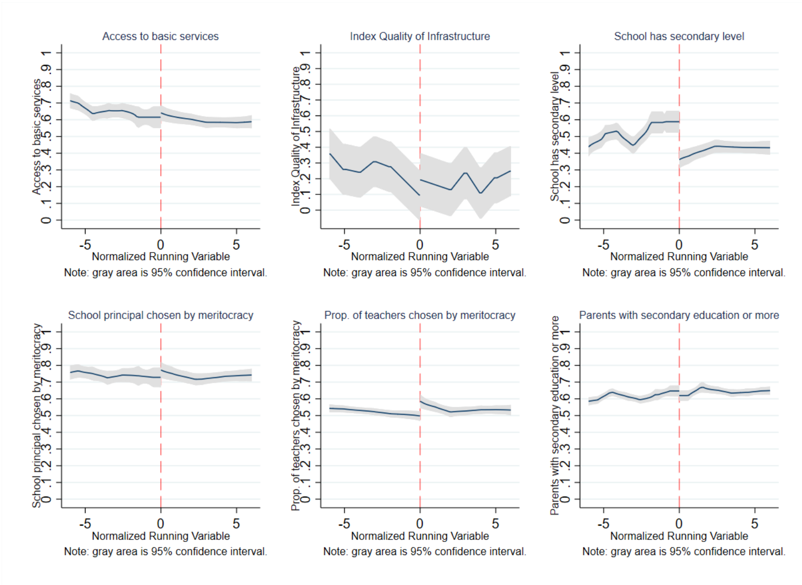


Figure 7: Continuity test of predetermined covariates



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 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%) LEMO it was not possible to implement the 2018 intervention and in around 10% of LEMO fewer activities were implemented due to logistic constraints[46]. Therefore, exposure to 2018 activities can be viewed as light-touch intervention.

MINEDU did not exclude schools exposed to 2018 intervention from the possibility of receiving

[46]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

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% of schools were assigned to treatment both in 2018 and 2019 and 15% 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 6, appendix 8.13). 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 8.13, I discuss the additional checks that I do to explore the effect of having in the sample few schools that were exposed to at least one activity of 2018 intervention.

Shortcomings of the empirical strategy: The use of a Fuzzy RDD allows to overcome threats to internal validity, but it also has 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)[47] and LATE estimates for all the schools that fall within the window of analysis of ±4. I observe that among eligible schools, the likelihood of reporting a case of violence increased by 37 percentage points[48]. The estimates are robust to the incorporation of covariates (columns 2 and 3) and to alternative the functional forms (columns 4 to 6). In column (3) and (7), I add a dummy variable that takes the value of 1 if the school was exposed to at least 1 activity of 2018 intervention. Consistent with the fact that eligibility to 2019 TA and 2018 intervention are independent, we observe that results remain almost unchanged after the incorporation of this covariate.

Among the schools that reported a case of violence, there are two types of schools: the schools that documented cases of school violence for the first time after the intervention and the schools that documented cases of violence before the intervention. Taking this into account, I estimate two separate regressions to explore which type of schools are driving the results. Table 2 - columns 1 and 2 - show the results from estimating equation 4.3 on an outcome variable that takes the value of 1 if the school registered cases of violence for the first time after the intervention, and 0 otherwise. Columns 3 and 4 show the results from estimating equation 4.3 on an outcome variable that takes the value of 1 if the school registered cases of violence before the intervention. The coefficient estimates are bigger

[47] The ITT are estimated using the following equation $y_{ij} = \delta_0 + \delta_1 T_{ij} + g(ranking_{ij}) + \gamma_j + \epsilon_{ij}$, 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.

[48] 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 8.10

among the first group of schools, suggesting that my estimates are driven by the schools that had not registered cases of violence before 2019.

Table 1: Likelihood of Reporting Violence

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV - LATE	0.461***	0.377***	0.372***	0.462***	0.377***	0.372***
	(0.110)	(0.127)	(0.126)	(0.110)	(0.127)	(0.126)
ITT	0.209***	0.141***	0.139***	0.209***	0.141***	0.139***
	(0.0516)	(0.0479)	(0.0475)	(0.0516)	(0.0479)	(0.0475)
F-stat	67.02	49.64	49.41	67.00	49.66	49.43
Anderson-Rubin Test P-values	0.0000695	0.00354	0.00379	0.0000694	0.00353	0.00378
N	1764	1755	1755	1764	1755	1755
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=2	p=2	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3). The first row of each Panel shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Columns (2) and (4) include school covariates, 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Likelihood of Reporting Violence for schools reporting for the first time before or after the intervention

	First-Time School Registered a Report		*School Registered Reports Before*	
	(1)	*(2)*	*(3)*	*(4)*
IV - LATE	0.230**	0.230**	0.138	0.138
	(0.102)	(0.102)	(0.0911)	(0.0911)
ITT	0.0855**	0.0855**	0.0513	0.0514
	(0.0388)	(0.0388)	(0.0363)	(0.0363)
F-stat	48.89	48.91	48.89	48.91
Anderson-Rubin Test P-values	0.0286	0.0286	0.159	0.158
N	1755	1755	1755	1755
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=2	p=1	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3). 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I also analyze the impact of the TA on the number of reports of school violence. Overall, I observe that in the schools that benefited from the TA, the number of reports of violence increased, on average, by 1 report[49]. To put this in context, as can be seen in figure 8, 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 (treated schools). Moreover, the mean reports of violence in the comparison group was 0.561, an average that is lower than 1 because many schools did not report any case of violence. Considering this, I keep the schools that reported cases of violence and observe that the mean and median number of reports of violence among the schools in the comparison group was 2 and 3, respectively. These descriptive statistics suggest that the average increase in 1 report of

[49]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 100%. This coincides with my findings from the ordinary least squares estimates (OLS) as the average increase in the reports of violence seems to be driven by first time reports, that move from 0 to 1 report of violence. See appendix 8.10.

violence might be capturing both an extensive margin increase that comes from the fact that a higher proportion of schools reported 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.

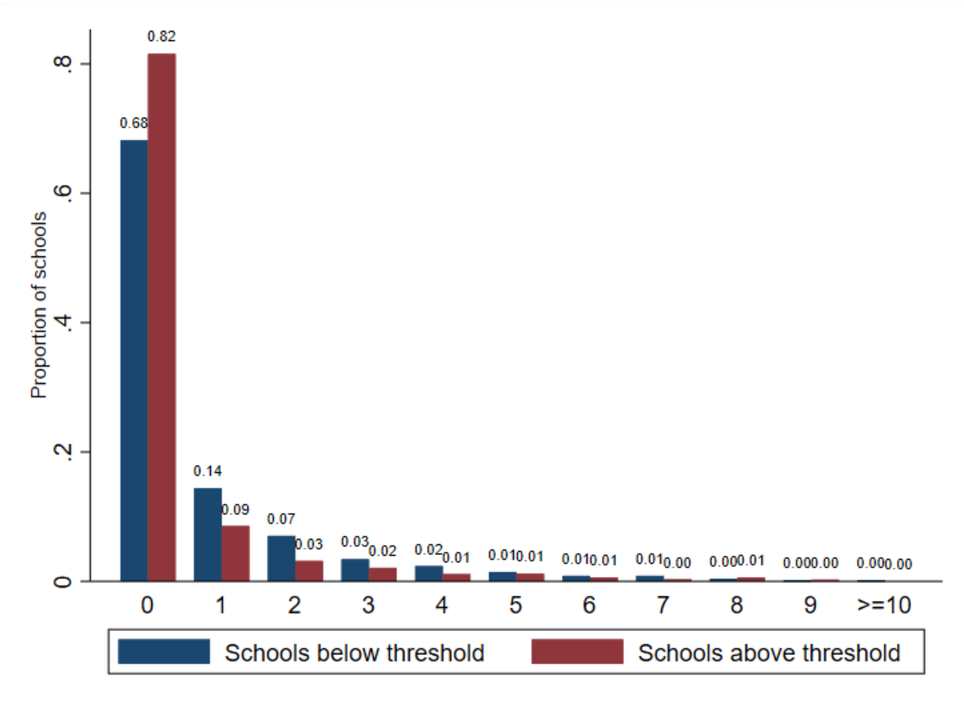
I am not able to analyze with precision whether the increase in the reports of violence is driven by cases of physical, psychological, or sexual violence. However, descriptive statistics show that in schools below and above the threshold the majority of reports of violence were cases of physical violence, followed-up by cases of psychological and sexual violence. The data also indicates that there are not statistically significant differences regarding the person that reported the incident of violence and type of perpetrator (see table 4, appendix 8.6). Importantly, the data does show that the increase in the number of reports of violence is driven mainly by secondary schools (see table 3, appendix 8.6).

Table 3: Number of Reports of School Violence

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV - LATE	1.157***	1.003**	0.995**	1.157***	1.003**	0.995**
	(0.380)	(0.434)	(0.434)	(0.380)	(0.434)	(0.434)
ITT	0.620***	0.373**	0.369**	0.620***	0.373**	0.369**
	(0.193)	(0.164)	(0.164)	(0.193)	(0.164)	(0.164)
F-stat	65.27	48.80	48.58	65.27	48.82	48.60
Anderson-Rubin Test P-values	0.00365	0.0243	0.0257	0.00365	0.0243	0.0257
N	1764	1755	1755	1764	1755	1755
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=2	p=2	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3).The first row of each Panel shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Columns (2) and (4) include school covariates, 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 8: Proportion of schools by the number of reports of violence



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 actions 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 that 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[50].

School, family, and individual factors will determine the weight individuals assign to the benefits and costs[51] 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 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 lead to a greater willingness to report events of violence.

A. Documenting cases of school violence for the first time

As discussed above (table 2), the schools that reported a case of violence for the first time after the intervention, seem to be driving my estimates of the overall likelihood of reporting violence. This is an interesting finding as these schools are probably the schools that faced the biggest barriers before 2019.

B. Working on school tasks that can reduce 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 learn that, before the intervention, not all 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 6 months after the intervention, indicates that more than two-thirds of the beneficiary schools work for the first time in tasks that could have reduced barriers to reporting. This includes practices related to the dissemination of the online reporting platform (SíSeVe), the creation of spaces for students to report and talk about school violence, the monitoring cases of school violence, and the execution of general meetings with the school community to talk about school violence[52].

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

[51]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).

[52]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 that responded to the

To explore this further, I also 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 by school principals and contained information about several school practices, including a few practices that were discussed through the TA[53]. I use this data to create an indicator of the total number of school practices related to the management of school violence and observe that the beneficiary schools implemented, on average, more practices. This finding provides additional supporting evidence about changes in school violence management practices that might have influenced the decision of reporting.

Table 4: Number of school practices related to the management of school violence

	(1)	*(2)*	*(3)*	*(4)*
IV-LATE	0.872***	0.797**	0.872***	0.797**
	(0.279)	(0.343)	(0.279)	(0.343)
F-stat	67.44	49.04	67.39	49.05
Anderson-Rubin Test P-values	0.00307	0.0252	0.00309	0.0253
N	1760	1752	1760	1752
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=2	p=2

Notes: The table presents coefficient estimates that represent the LATE and standard errors obtained after estimating our main specification (equation 4.3). Columns (2) and (4) include school covariates, and 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C. Student survey data on the presence of school violence

An alternative story could be that the increase in reporting is reflecting an increase in violence levels. This story is difficult to prove due to the issue of underreporting. 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% 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 being reported. For instance, in more than 90% of the schools that reported a case of violence for the first time in 2019, 2 out of 10 students had witnessed cases of physical and psychological violence[54] in 2018. 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.

Importantly, using this student-level dataset, I create an index of perceptions of school violence.

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, also suggests that 82% of the schools implemented changes in their school violence management practices

[53]The practices include: i) develop school coexistence plan for the school and classroom, ii) ensure access to an online reporting platform, iii) implement school violence prevention activities (workshops), iv) select a teacher representative responsible for school coexistence, among others.

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

This indicator is built using factor analysis and it is based on the student responses to several questions regarding whether they witnessed events of physical and psychological violence in the school. Table 5 shows no significant differences in the index of perception of school violence between the schools eligible to receive the TA and those that were not. This finding serves as suggestive evidence that there was not an increase in violence levels due to the intervention.

Table 5: Index of perception of school violence

	(1)	*(2)*	*(3)*	*(4)*
IV	-0.0902	-0.0932	-0.0639	-0.0623
	(0.169)	(0.157)	(0.177)	(0.164)
ITT	-0.0368	-0.0399	-0.0245	-0.0250
	(0.0656)	(0.0639)	(0.0657)	(0.0639)
F-stat	12.81	13.62	9.900	10.56
Anderson-Rubin Test	0.576	0.533	0.710	0.696
N	39163	38792	39163	38792
LEMO fixed effects	Yes	Yes	Yes	Yes
School & Individual Level Covariates	No	Yes	No	Yes
Exposure to TA 2018 Covariate	No	Yes	No	Yes
Polynomial	p=1	p=1	p=2	p=2

Notes: The outcome variable refers to an index of perception of school violence where higher values of the index indicate signal a higher prevalence of violence in the school. The table presents coefficient estimates that represent the LATE and standard errors obtained after estimating our main specification (equation 4.3). Columns (2) and (4) include school covariates, and 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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.

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), as well as with poorer student performance (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.

Table 6, Panel A and B, show the ITT and LATE on the likelihood of student dropout and student mobility, respectively. Columns (1) and (4) shows the estimates without including covariates, while column (2) and (5) include covariates at the school level and the individual level[55]. Importantly, the incorporation of the covariate that controls for exposure to 2018 intervention does not change the coefficient estimates (column 3 and 6).

The estimated coefficients suggest that the TA did not have statistically significant impacts on the likelihood of student dropout. 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

[55]Among the individual level covariates, I include a regressor on sex, age and parents' level of education. One caveat is that students' age does not vary smoothly around the threshold. However, estimates remain robust to the incorporation of the age covariate

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 carer's changes in employment or marital status - that lead to switching schools (Welsh, 2017), while the latter ones are related to the student's experiences in school. The TA is mainly likely to affect the decision of switching schools by improving the students' experiences in school. Therefore, to explore this further, I also account for residential mobility in the analysis.

The data allows me to observe if the student moved to schools located at the same district, or located at a different district, province or region[56]. 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[57]. Panel C, of Table 6, shows that the likelihood of student mobility decreased, on average, by 3 percentage points. The estimated coefficients are similar to those presented in Panel B, suggesting that the LATE are capturing mainly a reduction in school-related moves, rather than residential moves[58].

The potential channel explaining these results might be related to changes in school factors that have contributed to generating safe school spaces. As shown in Table 4, school principals of the beneficiary schools implemented more practices associated with the prevention and management of events of violence. Moreover, primary data collected from beneficiary schools also shows that two-thirds of the school principals implemented for the first time positive discipline strategies after the intervention. The latter involves avoiding corporal and verbal punishment, building a sense of community, and using effective communication to deal with misconduct. Therefore, even though this data is based on self-reports and might be noisy, it suggests that there has been a shift towards school practices that could have contributed to reducing the likelihood of switching schools through improving the overall school environment.

[56]Perú's territory is organized in regions, and these are subdivided into provinces that are composed of districts.

[57]In other words, I create a variable of non-structural mobility that takes the value only if the student move for reasons other than residential mobility, and zero otherwise.

[58]I also estimate a regression using as an outcome the indicator of residential mobility and observe that the intervention did not have an impact on this.

Table 6: IV Estimates: Dropout and Mobility

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Panel A: Dropout						
IV - LATE	0.00379	0.000758	-0.000257	0.00464	0.00171	0.000710
	(0.00754)	(0.00857)	(0.00855)	(0.00759)	(0.00857)	(0.00856)
ITT	0.00141	0.000247	-0.0000840	0.00174	0.000554	0.000231
	(0.00280)	(0.00280)	(0.00280)	(0.00281)	(0.00278)	(0.00279)
F-stat	17.29	15.21	15.33	17.15	14.65	14.76
Anderson-Rubin Test P-values	0.614	0.930	0.976	0.538	0.842	0.934
N	501092	497827	497827	501092	497827	497827
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=1	p=1	p=1	p=2	p=2	p=2
Panel B: Mobilitty						
IV - LATE	-0.0374**	-0.0344*	-0.0351**	-0.0365**	-0.0342*	-0.0349*
	(0.0178)	(0.0181)	(0.0177)	(0.0181)	(0.0187)	(0.0184)
ITT	-0.0139***	-0.0112**	-0.0115**	-0.0137***	-0.0111**	-0.0113**
	(0.00503)	(0.00471)	(0.00460)	(0.00514)	(0.00483)	(0.00471)
F-stat	17.29	15.21	15.33	17.15	14.65	14.76
Anderson-Rubin Test P-values	0.00609	0.0182	0.0133	0.00848	0.0227	0.0169
N	501092	497827	497827	501092	497827	497827
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=1	p=1	p=1	p=2	p=2	p=2
Panel C: Mobilitty (excluding residential mobility)						
IV - LATE	-0.0323**	-0.0289**	-0.0294**	-0.0317**	-0.0287**	-0.0291**
	(0.0135)	(0.0133)	(0.0132)	(0.0139)	(0.0140)	(0.0138)
ITT	-0.0120***	-0.00942***	-0.00960***	-0.0118***	-0.00929***	-0.00947***
	(0.00366)	(0.00341)	(0.00339)	(0.00378)	(0.00353)	(0.00349)
F-stat	17.29	15.21	15.33	17.15	14.65	14.76
Anderson-Rubin Test P-values	0.00117	0.00631	0.00502	0.00198	0.00904	0.00720
N	501092	497827	497827	501092	497827	497827
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=1	p=1	p=1	p=2	p=2	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3) for the window of ±4. The analysis is at the level of the students. Robust standard errors are reported in parenthesis and are clustered at the LEMO level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.3 Learning

To study the 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 7 summarizes the ITT and LATE of the TA on math and language test-scores. Even though that the estimated coefficients are positive, suggesting improvements in learning, these are not statistically significant. It is important to highlight that it might be too soon to detect impacts on learning. The national standardized tests were administered in November, and the intervention was implemented between May and October. Moreover, considering that my sample is reduced by 53% when only retaining secondary schools, it is also possible that I lack enough statistical power to detect impacts in these outcomes.

Table 7: IV Estimates: Learning

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Panel A: Math Scores						
IV - LATE	0.293	0.205	0.202	0.317	0.238	0.234
	(0.247)	(0.214)	(0.212)	(0.274)	(0.239)	(0.238)
ITT	0.118	0.0876	0.0862	0.119	0.0951	0.0937
	(0.0907)	(0.0862)	(0.0852)	(0.0932)	(0.0894)	(0.0886)
F-stat	12.56	13.90	13.75	9.646	10.76	10.65
Anderson-Rubin Test P-values	0.196	0.311	0.313	0.203	0.289	0.291
N	42438	40787	40787	42438	40787	40787
Panel B: Language Scores						
IV - LATE	0.300	0.156	0.150	0.126	0.0717	0.0694
	(0.244)	(0.187)	(0.185)	(0.0904)	(0.0779)	(0.0771)
ITT	0.120	0.0666	0.0642	0.335	0.179	0.174
	(0.0901)	(0.0763)	(0.0754)	(0.268)	(0.207)	(0.206)
F-stat	12.56	13.90	13.75	9.646	10.76	10.65
Anderson-Rubin Test	0.184	0.384	0.395	0.165	0.358	0.369
N	42452	40797	40797	42452	40797	40797
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=2	p=2	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3) for the window of ±4. The analysis is at the level of the students. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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 ± 4 window of analysis. In the appendix 8.7, table 4, 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 such as ± 3, 5 and 6. 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. We also have to keep in mind that in the ± 7 window, even though the estimated coefficients remain similar in size and statistical significance, the pre-treatment outcomes do not vary smoothly around the cut-off. Moreover, in the ± 8 and ± 9 window, the size of the estimated coefficient shrinks and the standard errors increase in size. Similarly, when analyzing the LATE on school mobility at different windows of analysis, I observe that the estimates only remain statistically significant and similar in magnitude in windows between ± 4 and ± 6 (see appendix 8.7, table 5).

Functional Form. I control for local linear and quadratic polynomials[59] and add interaction terms between the running variable and the treatment dummy, instrumenting for this with interactions between the running variable and the eligibility dummy (see appendix 8.8). The magnitude and direction of the estimated coefficients remain similar across all specifications, except in the model of quadratic polynomials with interactions 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 appendix 11.

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 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 8.10.

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

[59]Following Gelman and Imbens (2019), I do not consider higher order polynomials. The authors' advice 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 polynomial can produce confidence intervals that can lead to misleading inference.

alternative cut-off points. See appendix 8.12, figures 5a and 5b, where the grey line represents the true cut-off point, while the red line represents the placebo thresholds.

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 6, appendix 8.13). This confirms that treatment status in 2018 is independent to the eligibility in 2019. In addition to this, I estimate my main specification in two alternative samples: (i) a sample that drops the schools exposed to the intervention in 2018; and, (ii) a sample of schools located in the 40 LEMO that were only exposed to 2019 intervention. 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 8.13). This could be either because the results in school mobility were driven by the schools exposed also to the activities of 2018 intervention, or because I lose statistical power to detect impacts after dropping 22% of the schools in the sample that included 42% of the individual level observations.

7 Conclusion

School Violence is a worldwide phenomenon affecting almost a third of students aged 13 to 15 years (UNESCO, 2019). Extensive empirical evidence has shown the negative effects of being a victim, a bystander, and a perpetrator of school violence. However, there is still little rigorous evidence in low-middle income countries about the impact of interventions targeting the topic of school violence. This paper contributes to this research gap by exploring the effects of a Peruvian nationwide intervention that aimed to improve the management of school violence through training school heads 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 find that that the likelihood of reporting cases of violence increased by 37 percentage points, an effect that is mainly driven by the schools that reported cases of violence for the first time in 2019. I also analyse the number of reports of violence and observe that these increased by 1 report among the eligible schools.

Taken together, different pieces of evidence indicate that one plausible explanation for these results relates to a reduction in the barriers to reporting. First, the likelihood of reporting is mainly driven by the schools that might have faced the biggest barriers before the intervention: the first-time reporters. Second, qualitative and primary survey data, indicate that before the intervention, not all the school community knew how and where to report, and that there were fears and uncertainties related to reporting. A common statement from both the LEMO and school principals was that the intervention helped to overcome these constraints and that it contributed to normalizing the importance of reporting. Administrative data confirms this as the eligible schools worked more on tasks that could have reduced reporting barriers. Therefore, the estimates of this paper are more likely to inform about shifts in reporting behaviour rather than informing about actual changes in levels of violence. This finding is particularly relevant for policymakers as underreporting limits the possibility of dealing with and reducing future events of violence.

I also examine the impacts of the intervention on educational outcomes. The intervention had effects on reducing the likelihood of student mobility by 3 percentage points, effects that are comparable to other interventions that follow a student-only approach. I do not observe impacts in other outcomes including the dropout rate and student test scores.

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. 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 part 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 school violence.

References

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

Angrist, N., S. Djankov, P. K. Goldberg, and H. A. Patrinos (2021, April). Measuring human capital using global learning data. *Nature 592*(7854), 403–408.

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, August). 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, June). 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, November). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs: Robust Nonparametric Confidence Intervals. *Econometrica 82*(6), 2295–2326.

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

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

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

Delprato, M., K. Akyeampong, and M. Dunne (2017, January). 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, July). 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, February). 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.

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

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, June). 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, July). 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, September). 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.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2004, August). 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.

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

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, December). 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, August). 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, January). A Large-Scale Evaluation of the KiVa Antibullying Program: Grades 4-6: Evaluation of KiVa Antibullying Program. *Child Development 82*(1), 311–330.

Leaver, C., R. F. Lemos, and D. Dillenburg Scur (2019, November). 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, February). Regression discontinuity inference with specification error. *Journal of Econometrics 142*(2), 655–674.

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

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2020, October). Valid t-ratio Inference for IV. *arXiv:2010.05058 [econ, q-fin]*. arXiv: 2010.05058.

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

Limber, S. P., D. Olweus, W. Wang, M. Masiello, and K. Breivik (2018, August). 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.

Muralidharan, K. and A. Singh (2020, November). 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, November). 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, December). A useful evaluation design, and effects of the Olweus Bullying Prevention Program. *Psychology, Crime & Law 11*(4), 389–402.

Oreopoulos, P. (2007, December). 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, January). 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, January). Offender Decision-Making in Criminology: Contributions from Behavioral Economics. *Annual Review of Criminology 1*(1), 379–400.

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

Pritchett, L. (2013). *The Rebirth of Education: Schooling Ain't Learning.* Brookings Institution Press.

Rees, D., J. Sabia, and G. Kumpas (2020, February). Anti-Bullying Laws and Suicidal Behaviors among Teenagers. Technical Report w26777, National Bureau of Economic Research, Cambridge, MA.

Rivers, I., V. P. Poteat, N. Noret, and N. Ashurst (2009, December). 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, February). School Management, Grants, and Test Scores: Experimental Evidence from Mexico. Technical report, The World Bank.

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

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

Skogan, W. G. (1984, May). 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.

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, April). 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, February). 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.

Tomczyk, S., G. Schomerus, S. Stolzenburg, H. Muehlan, and S. Schmidt (2020, August). 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.

Varhama, L. M. and K. Björkqvist (2005, April). 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, June). 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. *24*(10), 1958–1970.

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

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

8 Appendix

8.1 Programs about Prevention of School Violence

Program	Countries	Program Approach	Description Intervention	Few references	Method	Findings: Violence-Related Ouctomes	Findings: Educational-Related Ouctomes	Findings: Other Ouctomes
Olweus Bully Prevention Program (OBPP)	Several countries inluding United States and Norway	Whole School Approach	The program includes several components targeting the school, the classroom, the individuals and even the community. *School-level components include:* Establish a Bullying Prevention Coordinating Committee (BPCC), Conduct trainings for the BPCC and all staff. Administer the Olweus Bullying Questionnaire (Grades 3-12), Hold staff discussion group meetings, Introduce the school rules against bullying, Review and refine the school's supervisory system, Hold a school-wide kick-off event to launch the program, Involve parents *Classroom-level components include:* Post and enforce school-wide rules against bullying, Hold regular (weekly) class meetings to discuss bullying and related topics, Hold class-level meetings with students' parents *Individual-level components include:* Supervise students' activities, Ensure that all staff intervene on the spot when bullying is observed, Meet with students involved in bullying, Meet with parents of involved students, Develop individual intervention plans for involved students, as needed *Community-level components include:* Involve community members on the BPCC, Develop school-community partnerships to support the school's program, Help to spread antibullying messages and principles of best practice inthe community.	Olweus (1993 & 1994)	ANOVA (time-comparison treated schools)	Reduction in bullying victimization rate. Higher peer-reported assisting.		
				Olweus and Limber (2010)	Literature rewview of papers studying the effect of OBPP.	Reduction in bullying victimization rate. Higher peer-reported assisting.		
				Limber et al (2011)	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	Classroom teachers carry out 20 hours of lessons involving discussion, role-play, video-clipsabout bullying, group work and written tasks (Kärnä et al., 2011a). Lessons focus onthe topic of bullying; children learn what bullying is, its different forms, consequencesand how individuals and groups can reduce it. The lessons also focus on socialskills; children learn about emotions, respecting others, being part of a team andgroup dynamics (Hutchings & Clarkson, 2015; Kärnä et al., 2013; Salmivali, Kärnä, &Poskiparta, 2010). Teachers: the teachers receive training and support to implement the program. Parents: the parents also receive information about bullying.	Kärnä et al. (2011)	Multilevel regression analysis (random assignment of schools to treated and control	Reduction in bullying victimization rate. Higher peer-reported assisting.		
				Nocentini and Menesini (2016)	Multilevel regression analysis (random assignment of schools to treated and control	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	Non curricular prevention strategy that follows the three-tiered prevention strategy. Tier 1 includes school wide components, while Tier 2 and Tier 3 are targeted to students at risk. Adoption of this program takes 2 to 3 years. Tier 1 activities involve defining, teaching, monitoring, and rewarding a small set of behavioral expectations for all students across non-classroom and classroom settings (Horner et al , 2009)	Bradshaw & Waasdorp (2015).	RCT. Latent Profile Analysis.	Lower likelihood of discipline deferrals.		
				Horner et al (2009)	Randomized, wait-list control effectiveness trial	Higher perceived safety of the school	Higher proportion of students achivieng state reading assessment	
				Flannery et al (2013)	Multilevel latent growth model	Lower likelihood of discipline deferrals.		

Program	Countries	Program Approach	Description Intervention	Few references	Method	Findings: Violence-Related Ouctomes	Findings: Educational-Related Ouctomes	Findings: Other Ouctomes
Student Success Through Prevention	United States	Student Only Approach	The intervention consisted on "15 weekly lessons of the sixth-grade curriculum that focused on social emotional learning skills, including empathy, communication, bully prevention, and problem-solving skills." (Espelage et al, 2013. Pg. 1) Sample of schools: 18 primary schools allocated to the treatment group and 18 to the control.	Espelage et al (2013)	RCT	Deecrease in the likelihood of self-reporting physical aggression	-	-
Raising Voices Uganda - The Good School Toolkit	Uganda	Whole School Approach	The intervention is a violence prevention behavioural intervention called The Good School Toolkit. The intervention included around 60 activities directed to the staff, students and administration 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 schools goals. Sample of schools: 21 primary schools allocated to the treatment group and 21 to the control.	Knight et al (2018) Devries et al (2015)	RCT	Decrease in the likelihood of experiencing physical violence from staff	-	-
Guitierrez et al (2008)	Peru	Student only approach	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. Sample of schools: 33 secondary public schools were randomly allocated to treatment and 33 to the control group.	Gutierrez et al (2018)	RCT	Increase in the likelihood of reporting school violence.	Reduction in the likelihood of dropout and mobility. Increase in test-scores.	Decrease in depression.
AntibuLllying LAWS	United States	-	AntiBullying Laws (ABLs) include laws were categorized in two groups: strong and weak. Strong ABLs include at least 3 of 5 requirements, whether weak laws included 0 2. The requirements were: "school districts must (i) provide written records of bullying and how each incident was 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 behaviors that constitute bullying." Ress et al, 2020. Pg 14.	Ress et al 2020	Differences in Differences	Decrease bullying victimization	-	Decrease depression and suicidal ideation
Skill Based Violence Prevention Program (VPP)	United States	Student only approach	The intervention has 12 sessions. The sessions aim is ti build skills ti improve conflict-related attitudes and behaviours. The topics of the sessions include self-concept, group dynamics, vision and imagination, conflict management and communication skills. The sample was only composed by 13 schools.	Thompkins et al (2013)	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
			The intervention is a cognitive behaviour program that consists of 10 sessions. The topics of the sessions include discussion about the problem of violence, understanding feelings and stimuli, discriminating between thoughts and emotions, and building problem solving skills, among others. and identidying changes in self-concept, group dynamics, vision and imagination, conflict management and communication skills. The sample is composed by 1 school (157 students).	De Anda (1999)	T-tests - before and after (only treated students).	Increase in the sense of safety in the school. Decrease in the acceptance of violence. Increase in the use of non-violent means of conflict resolution.	-	-

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).

8.2 Measurement: Outcome Variables and Covariates

Variable	Definition	Level	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.	School Level	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.	Report Level / School	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 Level / School	School 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 Level / School	School Census
Math and Language Test-scores	Math and language test-scores from the ECE. Scores are estimated using the Rasch Model.	Student Level / School	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.	School Level	List of targeted schools
Eligibility Dummy	Dummy variable takes the value of 1 if the school is located below the threshold rule of 9, and zero otherwise.	School Level	School Census
Running Variable	Discrete running variables normalized to zero. The minimum value is 9 and the max value depends on the number of schools located in a specific LEMO. Each value indicates the position of the school in a ranking that is based on the minimun distance of the adjacent school to the nucleo school, and the number of enrolled students.	School Level	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 Level	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 Level	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 Level	School Census
Proportion of teachers chosen by meritocracy		School Level	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.	Student Level / School Level	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 Level	School Census
Student sex	Dummy variable takes the value of 1 if the student is a male, and zero otherwise.	Student Level	School Census
Student age	Student age.	Student Level	School Census
Other			
Changes in School Practices after the intervention	The survey asked about 8 school practices: Design School Coexistence Rules, Monitor the reported cases of violence, Dissemination of the available channels of reporting, Create safe spaces of dialogue within the school premises, Promote the use of positive discipline strategies, Hold general meetings with all the school members to discuss the topic of school violence, Hold meetings with external specialized agencies that provide in the prevention and management of cases of violence. The survey also asks whether the school had work on these activities before the intervention or not. If they had, it asked about changes in frequency of time dedicated to these activities. With this information I create an index of changes in school practices, as well as alternative variables related to the Total number of new practices implemented.	School Level	School Survey
Time committed to tasks suggested in the intervention	Dummy variable that takes the value of 1 if the	School Level	School Survey
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 Level	School Census

8.3 LEMO and School Survey

School survey data: the school principals of the 2,650 beneficiary schools (1852 schools with primary and secondary level and 798 schools with only primary level) were invited to respond an online school survey between July 1st and September 1st of 2020, 6 months after 2019 intervention[60]. In addition to the email invitation to respond the survey, all survey participants were contacted by phone to inform them about the survey objectives and to offer the option of responding 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 internet. In total, 1,235 schools responded the survey: 54% of secondary schools with primary and secondary level, and 29% of primary schools. 59 School Principals refuse to answer, 83 initiated the survey but completed less than 50% of the survey and the remaining did not respond the survey, had phone numbers that were not answered, were non-existent or went to voicemail.

The School Survey data is mainly representative for urban schools. 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, scales of alignment and the Facilita-

[60]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 of 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.

tors quality, time committed to the activities proposed in the Technical Assistance, practices or tasks performed before and after the intervention, scale of knowledge of the topics of the Technical Assistance, perceptions of the main impacts of the Technical Assistance, among others. This paper mainly uses information about the time committed to the activities proposed in the Technical Assistance and the practices or tasks performed before and after the intervention.

LEMO survey data: all LEMO Facilitators were invited to respond an online survey between April 8th and June 8th of 2020, 4 months after 2019 intervention. The LEMO Facilitators that worked in 176 local offices were successfully interviewed, resulting in a respond rate of 80%. This can be considered a significant response rate considering that the survey was voluntary, it lasted around 40 minutes and was implemented during the lockdown months imposed in Peru due to the COVID-19 pandemic. Considering that I do not have survey data for 44 local offices , I test the null of no differences in means between those that responded the survey and those that didn't, and I estimate a logistic regression model to analyse if there are any covariates that are correlated with the likelihood of responding the survey. Results show that the covariates are uncorrelated with the likelihood of responding the survey, except for the dummy variable measuring whether the Facilitator was still working in the LEMO in 2020. This coefficient is positive and statistically significant at the 10 percent level, indicating that those that responded the survey are more likely to be those that were working in the local office in 2020. This finding is not surprising as some survey participants were not sure if they could answer the survey considering that they were not currently working in the local office, and that others specified they were not interested in responding the survey because they were not working in the local office anymore. We might worry that those that were not working in 2020 and did not respond the survey, were also the ones that performed worst. Yet, the data shows that this is not the case. Only 24% of those that were not hired in 2020 did not respond the survey and there is not a statistically significant difference in the proportion of programme components delivered among those that were not hired in 2020 and completed the survey and those that did not complete the survey.

The survey collected data on the following variables: individual characteristics of the Facilitator, characteristics of the LEMO (e.g., management index), delivery of the Technical Assistance, scales of the perception of alignment of the school and the Facilitators quality, among others.This data is mainly used in a companion paper.

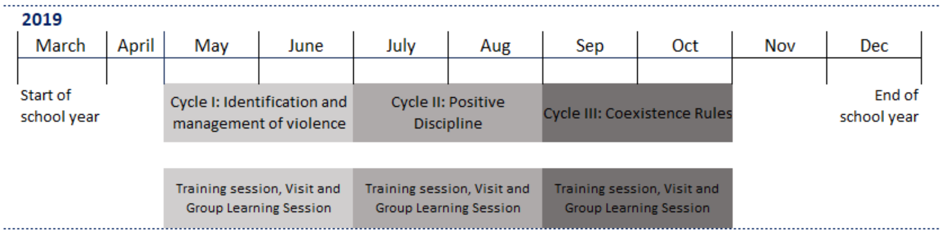
8.4 Curriculum of the Technical Assistance

The Technical Assistance was structure in three cycles[61] (figure 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 on 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

[61] 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

group learning sessions were executed among sub-groups of 4 targeted schools (1 nucleo school and 3 adjacent schools)[62]. During these sessions, the schools shared their experiences working on the topic that was discussed during the training session.

Figure 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[63]; iv) how to use the online platform SíseVe to monitor and manage the cases of violence registered by students, confidants of the victim and bystanders. The training also highlighted that the importance of informing about the reporting platforms to all the school community.

- *Second Cycle: Positive Discipline..* The training topics included: i) challenges on 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 school, and for each classrooms. The training highlighted that the coexistence rules had to be develop with all the school community and that they had to be published in a visible location within the school and the classroom.

8.5 Methodology to estimate the Running Variable

In 2019, with the aim of reaching schools with different situations of school violence, the Ministry targeted two type of schools:

- Nucleo schools: in each LEMO, the Ministry selected 3 schools with the highest incidence of violence, highest number of enrolled students and lowest travel time distance to the LEMO. These were the schools were the group sessions were executed.

- Adjacent schools: in each LEMO, the Ministry selected 9 schools. The main criteria to select these schools was based on the distance to the Nucleo schools. For each Nucleo school, the Ministry mapped all the schools that were in each LEMO and selected the 3 schools that were closest in distance to each Nucleo school. When two schools were at the same or similar distances,

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

[63]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.

they choose the school that had the higher number of enrolled students. This eligibility criteria were made to minimize the risks of low attendance to the group sessions that could exist due to travel time costs. Moreover, it allows to target schools that were more similar to the median public school, as well as schools that might have been underreporting or not reporting cases of violence.

The eligibility criteria that was used to select the Adjacent schools provides an opportunity to study the impact of the TA in these schools. First, the selection of the schools was based on two variables that aren't possible to manipulate and that I assume to be independent of potential outcomes. Second, the random cut-off of 9 provides an opportunity to compare the schools that were just below and just above the cut-off rule, as I expect that those schools above the cut-off, weren't selected just because they were a few kilometres farther away to the nearest Nucleo school and/or because they had fewer number of enrolled students. Therefore, to exploit the exogenous variation created by the eligibility criteria, I will try to replicate the selection process followed by the Ministry.

To select the Adjacent schools, the Ministry of Education mapped all the public schools in the country and, for each of the 221 LEMO, selected the beneficiary schools by looking at the map and selecting the 3 schools that were closest to each Nucleo school (a total of 9 schools per LEMO). The selection of schools was based on eyeballing which schools were closest to the Nucleo school and, as a result, the Ministry didn't record information about the distance from the public schools to each Nucleo school nor which schools were just few kilometres farther away from each Nucleo school. Therefore, I proceed to replicate the eligibility criteria by creating a ranking of schools based on their distance to each Nucleo school and the number of enrolled students. Below I explain in detail the steps that were followed:

1. Estimate the distance to each Nucleo school: the distances are calculated using Vicenty (1975) formula. The method calculates the distances between a pair of latitude and longitude points assuming an oblate sphere or an ellipsoidal model of the Earth.[64]

2. Rank the schools based on their distance to each Nucleo school: in 12% cases [65], schools can 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 to other Nucleo schools. [66]

   Step 2 generates three rankings of distance per LEMO that are associated to each Nucleo school. Each Adjacent school in the ranking will have values between 1 and Z, where Z represents the school that is further away to the Nucleo school.

3. Rank the schools based on the population or number of enrolled students: I ranked the schools

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

[65] 12% 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%

[66] Each LEMO has 3 Nucleo schools. 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 - Nucleo 1- , then I estimate the ranking of distance in the subsequent Nucleo - 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 - 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. For the estimation of the ranking of distance in each Nucleo, I start with either Nucleo 1, 2 or 3 depending on which ranking estimation in each Nucleo and in each LEMO has the highest rate of predictability relative to the true assignment dummy provided by the Ministry.

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.

4. Assign weights to the *distance* and *population of students* ranking: considering that the distance and the population of students were used in the process of assigning Adjacent schools to treatment, I analyse the importance of each variable in explaining the eligibility criteria. Based on Slottje (1991), Boysen (2002) and Poppitz (2019), I estimate equation 8.1 using an Ordinary Least Square model where the dependent variable is a dummy that takes the value of 1 for schools that received the TA and use the regression coefficients to create weights [67] for the two independent variables used in the model. This analysis suggests that the weights that have to be used are, in average, $W_{distance}$=0.90 and $W_{population}$=0.10.

$$T_{ij} = \beta_0 + \beta_1 Distance_{ij} + \beta_2 PopulationStudents_{ij} + \epsilon_{ij} \tag{8.1}$$

Qualitative interviews with the civil servants that designed the intervention reveal that the importance given to the distance and population variable varied between LEMO, mainly depending on the density of schools by LEMO. Considering this, I explore 11 different weighting schemes ranging between $W_{distance}$=1 and $W_{population}$=0 and $W_{distance}$=0.60 and $W_{population}$=0.40[68]. This means that for each school I create a score that is associated to each of the 11 weighting schemes following equation 8.2.[69]

$$\begin{gathered} Score_{ijw} = RankDistance_{ij} W^{w}_{distance} + RankPopulation_{ij} W^{w}_{population}, \\ \text{where i=1 to N school, j=1 to 221 LEMO and w=1 to 11 weighting schemes} \end{gathered} \tag{8.2}$$

5. Rank schools based on the *distance* and *population of students* weighted ranking: I calculate a score for each school using equation 4.1 for each of the 11 weighting schemes. I then rank schools on ascending order based on the score and normalize the created ranking to zero. This indicator represents the running variable. I also create a dummy variable that represents the predicted treatment variable and takes the value of 1 if the value of the running variable is below zero and 0 otherwise. For each LEMO, I use the weighting scheme that yields the highest predictibility rate or the higher proportion of schools assign to treatment based both on the official treatment dummy provided by the Ministry and the predicted treatment dummy.

[67] The weights can be formally described by $W_{distance} = \frac{\beta_1}{\beta_1+\beta_2}$ and $W_{population} = \frac{\beta_2}{\beta_1+\beta_2}$

[68] I do not use weights lower than 0.4 because i) when constructing weights using regression analysis the results indicate that the average weight of the distance variable was 0.9 and ii) when repeating the latter exercise for each LEMO separately, weights below 0.4 for the distance variable were unlikely.

[69] Considering that for each school I have 3 rankings of distance, equation 8.2 is estimated for all the weighting schemes for each of the 3 distance rankings (33 times). After doing these, for each weighting scheme I chose the combination of distance ranking and population that gives the min score. This procedure ensures that each Nucleo School is allocated 3 Adjacent Schools.

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

8.6 Tables and Figures

Figure 2: Students' perceptions of school violence

(a) Violence Between students

(b) Violence teacher to student

Figure 3: Density of Schools Around the Threshold

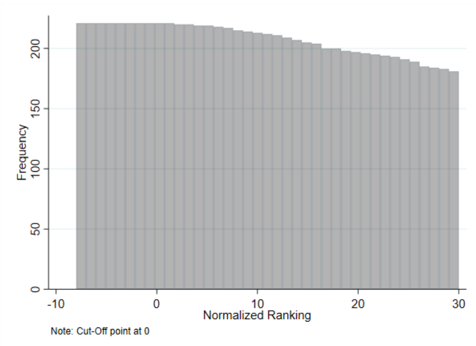


Table 1: Normalized Differences

	Treatment Mean	Control Mean	Normalized Difference
A. Pre-treatment outcomes			
Number of Reports of Violence	0.312	0.217	0.084
	(1.209)	(1.044)	
First Time Reported Violence	0.057	0.040	0.078
	(0.231)	(0.196)	
Proportion of students that dropout	0.030	0.029	0.029
	(0.034)	(0.033)	
Proportion of students that transferred to another school	0.049	0.050	0.032
	(0.049)	(0.047)	
B. Covariates			
School has access to electricity, water and sanitation	0.635	0.609	0.053
	(0.482)	(0.488)	
Index of School Infrastructure (material of construction)	0.229	0.167	0.049
	(1.218)	(1.276)	
School Principal chosen by Meritocracy	0.730	0.735	0.013
	(0.444)	(0.441)	
Proportion of teachers chosen by meritocracy	0.509	0.543	0.104
	(0.313)	(0.324)	
Proportion of parents with secondary education or more	0.625	0.639	0.060
	(0.246)	(0.238)	
School Has Secondary Level	0.5317	0.416	0.233
	(0.499)	(0.493)	

Notes: The table shows the normalized differences between schools located below and above the threshold in the window of analysis of ± 4. Normalized Differences larger than 0.25 indicate that the average covariate values are different between the two groups.

Table 2: Placebo Estimates

	N	Control Mean	Placebo
A. Pre-treatment outcomes			
School Reported Violence	1764	0.09	0.05
		(0.286)	(0.072)
Number of Reports of Violence	1764	0.217	0.313
		(1.044)	(0.242)
First Time Reported Violence	1764	0.040	0.020
		(0.196)	(0.051)
Proportion of students that dropout	1764	0.029	-0.000
		(0.033)	(0.008)
Proportion of students that transferred to another school	1764	0.050	-0.018
		(0.047)	(0.012)
B. Covariates			
School has access to electricity, water and sanitation	1764	0.609	-0.096
		(0.488)	(0.106)
Index of School Infrastructure (material of construction)	1763	0.167	-0.190
		(1.276)	(0.267)
School Principal chosen by Meritocracy	1764	0.735	-0.066
		(0.441)	(0.100)
Proportion of teachers chosen by meritocracy	1756	0.543	-0.194***
		(0.324)	(0.074)
Proportion of parents with secondary education or more	1764	0.639	-0.018
		(0.238)	(0.034)
School Has Secondary Level	1764	0.416	0.592***
		(0.493)	(0.115)

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 our main specification (equation 4.3) but using as dependent variable the outcomes and relevant covariates determined prior to the intervention in the year 2017 (results remain consistent when using 2018 as the pre-treatment year). The estimates correspond to a window of analysis of ± 4. Controls excluded from the table include quadratic distance to cutoff. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 4: Descriptive statistics: Type of Reports

	Type of violence			Main perpetrator		Who reports	
	Physical	Psychological	Sexual	School Staff	Students	School Staff	Other
Treated Group	48.50%	35.89%	15.62%	45.78%	54.22%	45.55%	54.45%
Comparison Group	49.27%	35.91%	14.82%	47.70%	52.30%	50.32%	49.68%
p-value	*0.80*	*0.99*	*0.71*	*0.52*		*0.21*	

Table 3: Violence related outcomes by grade

	Primary Level	Secondary Level	All	Primary Level	Secondary Level	All
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
IV: Likelihood of Reporting Violence	0.253**	0.680**	0.372***	0.242**	0.732**	0.372***
	(0.123)	(0.321)	(0.126)	(0.123)	(0.371)	(0.126)
IV: Number of Reports of Violence	0.565*	2.460*	0.995**	0.523	2.495	0.369**
	(0.328)	(1.403)	(0.434)	(0.328)	(1.608)	(0.164)
F-stat						
Anderson-Rubin Test						
N	1386	828	1755	1386	828	1755
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=2	p=2	p=2

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

8.7 IV Estimates in Alternative Windows

Table 4: IV Estimates in different Windows: Violence Related Outcomes

Window of Analysis	N	Likelihood of Reporting Violence (1)	Likelihood of Reporting Violence (2)	Number of Reports of Violence (1)	Number of Reports of Violence (2)
2	879	0.203	0.200	0.400	0.401
		(0.420)	(0.420)	(1.120)	(1.119)
3	1319	0.305	0.310	1.228*	1.242*
		(0.207)	(0.207)	(0.717)	(0.717)
4	1755	0.372***	0.372***	0.995**	0.995**
		(0.126)	(0.126)	(0.434)	(0.434)
5	2192	0.282**	0.282**	0.995**	0.994**
		(0.114)	(0.114)	(0.447)	(0.447)
6	2629	0.245**	0.245**	0.719*	0.719*
		(0.110)	(0.110)	(0.430)	(0.430)
7	3063	0.290***	0.290***	0.649	0.650
		(0.0974)	(0.0974)	(0.407)	(0.407)
8	3496	0.265***	0.266***	0.383	0.384
		(0.0986)	(0.0986)	(0.411)	(0.411)
9	3927	0.190**	0.189**	0.406	0.406
		(0.0840)	(0.0840)	(0.340)	(0.340)
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=2	p=1	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (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 5: IV Estimates in different Windows: Likelihood of School Mobility

Window of Analysis	N	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Mobility									
2	246664	-0.0749	-0.0704	-0.0785	-0.0928	-0.0819	-0.0803	-0.0813	-0.104
		(0.109)	(0.0834)	-0.081	-0.11	(0.118)	(0.0994)	-0.077	-0.0486
3	372111	-0.0414	-0.0462	-0.04	-0.0449	-0.0450	-0.0503	-0.0433	-0.0486
		(0.0300)	(0.0349)	-0.0268	-0.0365	(0.0328)	(0.0384)	-0.0286	-0.0389
4	501092	-0.0379**	-0.0391**	-0.0374**	-0.0351**	-0.0369**	-0.0389**	-0.0365**	-0.0349*
		(0.0176)	(0.0183)	-0.0178	-0.0177	(0.0177)	(0.0189)	-0.0181	-0.0184
5	639136	-0.0287*	-0.0292	-0.0243*	-0.02	-0.0282*	-0.0304	-0.0238	-0.0209
		(0.0168)	(0.0180)	-0.0146	-0.0153	(0.0169)	(0.0189)	-0.0147	-0.0158
6	776256	-0.0262**	-0.0282**	-0.0231**	-0.0208*	-0.0271**	-0.0298*	-0.0238*	-0.0221
		(0.0122)	(0.0137)	-0.0112	-0.012	(0.0132)	(0.0154)	-0.0122	-0.0135
7	913014	-0.0200*	-0.0191	-0.016	-0.0126	-0.0196	-0.0192	-0.0158	-0.013
		(0.0117)	(0.0125)	-0.0108	-0.0115	(0.0125)	(0.0138)	-0.0116	-0.0127
8	1049847	-0.0169	-0.0175	-0.014	-0.0119	-0.0173	-0.0186	-0.0146	-0.013
		(0.0120)	(0.0128)	-0.0112	-0.0121	(0.0129)	(0.0141)	-0.0119	-0.013
9	1201534	-0.0196	-0.0219	-0.0176	-0.0169	-0.0215*	-0.0241	-0.019	-0.0185
		(0.0124)	(0.0145)	-0.0115	-0.0132	(0.0129)	(0.0153)	-0.012	-0.0138
Panel B: Mobility (excluding residential mobility)									
2	246664	-0.0569	-0.0475	-0.0723	-0.0708	-0.0647	-0.0569	-0.199	-0.0796
		(0.0810)	(0.0591)	-0.0734	-0.0868	(0.0910)	(0.0719)	-0.416	-0.0938
3	372111	-0.0362	-0.0370	-0.0370*	-0.0374	-0.0388	-0.0399	-0.0739	-0.0395
		(0.0225)	(0.0257)	-0.0216	-0.0294	(0.0247)	(0.0283)	-0.0661	-0.0308
4	501092	-0.0340**	-0.0343**	-0.0318**	-0.0294**	-0.0329**	-0.0337**	-0.0490**	-0.0287**
		(0.0132)	(0.0133)	-0.0128	-0.0132	(0.0133)	(0.0137)	-0.0224	-0.014
5	639136	-0.0280**	-0.0288**	-0.0249**	-0.0193	-0.0278**	-0.0300**	-0.0291	-0.0199
		(0.0133)	(0.0145)	-0.0115	-0.0122	(0.0135)	(0.0152)	-0.0187	-0.013
6	776256	-0.0219**	-0.0233**	-0.0211**	-0.0164*	-0.0239**	-0.0261**	-0.0202*	-0.0185*
		(0.00943)	(0.0107)	-0.00893	-0.00958	(0.0104)	(0.0123)	-0.0115	-0.011
7	913014	-0.0138	-0.0131	-0.0134	-0.00604	-0.0144	-0.0141	-0.00941	-0.00692
		(0.00875)	(0.00945)	-0.00874	-0.00888	(0.00958)	(0.0107)	-0.0116	-0.0102
8	1049847	-0.0113	-0.0117	-0.0107	-0.00384	-0.0131	-0.0138	-0.00658	-0.00535
		(0.00892)	(0.00965)	-0.00873	-0.00933	(0.00979)	(0.0108)	-0.0135	-0.0104
9	1201534	-0.0140	-0.0158	-0.0142	-0.00949	-0.0166*	-0.0184	-0.0142	-0.0111
		(0.00925)	(0.0109)	-0.00887	-0.0103	(0.00969)	(0.0115)	-0.0137	-0.0109
School Covariates		No	Yes	No	Yes	No	Yes	No	Yes
Exposure to 2018 intervention covariate		No	Yes	No	Yes	No	Yes	No	Yes
LEMO fixed effects		No	No	Yes	Yes	No	No	Yes	Yes
Polynomial		p=1	p=1	p=1	p=1	p=2	p=2	p=2	p=2

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

8.8 IV Estimates with Different Functional Forms

Table 6: IV Estimates: Reports of School Violence

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Panel A: Violence Outcomes								
IV: Likelihood of Reporting Violence	0.461***	0.372***	0.452***	0.363***	0.462***	0.372***	0.714***	0.460*
	(0.110)	(0.126)	(0.113)	(0.127)	(0.110)	(0.126)	(0.158)	(0.247)
F-stat	67.02	49.41	13.22	18.36	67.00	49.43	5.014	0.521
Anderson-Rubin Test P-values	0.0000695	0.00379	0.000259	0.0129	0.0000	0.00378	0.0000	0.00140
IV: Number of Reports of Violence								
	1.157***	0.995**	1.152***	0.990**	1.157***	0.995**	1.432	0.876
	(0.380)	(0.434)	(0.380)	(0.433)	(0.380)	(0.434)	(2.523)	(0.858)
F-stat	65.27	48.58	14.88	19.56	65.27	48.60	0.0266	0.451
Anderson-Rubin Test P-values	0.00365	0.0257	0.0144	0.0823	0.00365	0.0257	0.0251	0.134
N	1764	1755	1764	1755	1764	1755	1764	1755
Panel B: Dropout and Mobility Outcomes								
IV: Dropout	0.00379	-0.000257	0.00435	0.000508	0.00464	0.000710	0.0324	0.149
	(0.00754)	(0.00855)	(0.00770)	(0.00885)	(0.00759)	(0.00856)	(0.0923)	(2.658)
F-stat	17.29	15.33	3.345	3.838	17.15	14.65	0.0328	0.00104
Anderson-Rubin Test P-values	0.614	0.976	0.476	0.507	0.538	0.842	0.781	0.885
IV: Mobility (excluding residential mobility)	-0.0323**	-0.0294**	-0.0321**	-0.0294**	-0.0317**	-0.0291**	-0.0211	-0.0190
	(0.0135)	(0.0132)	(0.0137)	(0.0137)	(0.0139)	(0.0138)	(0.0485)	(0.318)
F-stat	17.29	15.33	3.345	3.838	17.15	14.76	0.0328	0.00104
Anderson-Rubin Test P-values	0.00117	0.00502	0.00400	0.0176	0.00198	0.00720	0.00348	0.0160
N	501092	497827	501092	497827	501092	497827	501092	497827
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School & Individual Level Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Exposure to 2018 intervention covariate	No	Yes	No	Yes	No	Yes	No	Yes
Interactions: eligible dummy&running variable	No	No	Yes	Yes	No	No	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=2	p=2	p=2	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3) for the window of ±4. Robust standard erros are reported in parenthesis and are clustered at the LEMO level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8.9 Alternative estimations of standard errors

Table 7: IV Estimates: Reports of School Violence

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Likelihood of Reporting Violence	0.461***	0.461***	0.372***	0.372***	0.372***	0.372***
	(0.110)	(0.0633)	(0.126)	(0.0756)	(0.126)	(0.0717)
Number of Reports of Violence	1.157***	1.157***	1.003**	1.003***	0.995**	0.995***
	(0.380)	(0.159)	(0.434)	(0.167)	(0.434)	(0.160)
Dropout	0.00379	0.00379	-0.000257	-0.000257	0.000710	0.000710
	(0.00754)	(0.00467)	(0.00855)	(0.00489)	(0.00856)	(0.00329)
Mobility	-0.0374**	-0.0374***	-0.0351**	-0.0351***	-0.0349*	-0.0349***
	(0.0178)	(0.00773)	(0.0177)	(0.00938)	(0.0184)	(0.0103)
Mobility (excluding residential mobility)	-0.0323**	-0.0323***	-0.0294**	-0.0294***	-0.0291**	-0.0291***
	(0.0135)	(0.00579)	(0.0132)	(0.00748)	(0.0138)	(0.00860)
N	1764	1764	1755	1755	1755	1755
School Covariates	No	No	Yes	Yes	Yes	Yes
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Exposure to 2018 intervention covariate	No	No	Yes	Yes	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=2	p=2
SE Clustered by	LEMO	Running Variable	LEMO	Running Variable	LEMO	Running Variable

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3) for the window of ±4. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8.10 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 8.3 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 8.3 will gives the reduced-form intent-to-treat estimates.

$$Y_{ij} = exp(\beta + \lambda T_{ij} + g(ranking_j)) + \mu_{ij} \tag{8.3}$$

  Table 8 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%.

Table 8: Reports of School Violence

	(1)	*(2)*	*(4)*	*(5)*
ITT - OLS: Number of Reports of Violence	0.620***	0.369**	0.620***	0.369**
	(0.193)	(0.164)	(0.193)	(0.164)
ITT - POISSON: Number of Reports of Violence	0.956***	0.862***	1.015***	0.878***
	(0.281)	(0.274)	(0.310)	(0.289)
N	1764	1755	1764	1755
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=2	p=2

Notes: The table presents the estimated coefficients for the window of ±4. 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Binary dependent variables

  Table 9 and 10, Panel A, presents the 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 different estimation models. Regarding the likelihood of mobility we observe similar ITT estimates. However, when using a bivariate probit, we observe that the size and statistical significance goes away after controlling for fixed effects and covariates.

Table 9: Likelihood of reporting violence

Panel A: ITT - OLS and Logistic Model	*(1)*	*(2)*	*(4)*	*(5)*
OLS: Likelihood of reporting violence	0.209***	0.139***	0.209***	0.139***
	(0.0516)	(0.0475)	(0.0516)	(0.0475)
LOGIT: Likelihood of reporting violence	0.278***	0.172***	0.283***	0.178***
	(0.0632)	(0.0574)	(0.0645)	(0.0584)
Panel B: LATE - IV and BiProbit Model	*(1)*	*(2)*	*(4)*	*(5)*
IV: Likelihood of reporting violence	0.461***	0.372***	0.462***	0.372***
	(0.110)	(0.126)	(0.110)	(0.126)
BiProbit: Likelihood of reporting violence	0.375***	0.320**	0.379***	0.334**
	(0.046)	(0.138)	(0.045)	(0.164)
N	1764	1755	1764	1755
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=2	p=2

Notes: The table presents the estimated coefficients for the window of ±4. 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Likelihood of reporting violence

Panel A: ITT - OLS and Logistic Model	*(1)*	*(2)*	*(4)*	*(5)*	*(5)*	*(6)*	*(7)*	*(8)*
Likelihood of school mobility								
OLS	-0.0141***	-0.0130***	-0.0139***	-0.0115**	-0.0138***	-0.0129***	-0.0137***	-0.0113**
	(0.00510)	(0.00477)	(0.00503)	(0.00460)	(0.00520)	(0.00487)	(0.00514)	(0.00471)
LOGIT	-0.0141***	-0.0127***	-0.0141***	-0.0116***	-0.0138***	-0.0126***	-0.0139***	-0.0115***
	(0.00507)	(0.00461)	(0.00494)	(0.00438)	(0.00515)	(0.00468)	(0.00503)	(0.00445)
Likelihood of school mobility (excluding residential mobility)								
OLS	-0.0127***	-0.0114***	-0.0120***	-0.00960***	-0.0123***	-0.0112***	-0.0118***	-0.00947***
	(0.00368)	(0.00346)	(0.00366)	(0.00339)	(0.00376)	(0.00351)	(0.00378)	(0.00349)
LOGIT	-0.0126***	-0.0110***	-0.0120***	-0.00924***	-0.0124***	-0.0108***	-0.0119***	-0.00916***
	(0.00366)	(0.00337)	(0.00358)	(0.00323)	(0.00373)	(0.00339)	(0.00367)	(0.00328)
Panel B: LATE - IV and BiProbit Model								
Likelihood of school mobility								
OLS	-0.0379**	-0.0391**	-0.0374**	-0.0351**	-0.0369**	-0.0389**	-0.0365**	-0.0349*
	(0.0176)	(0.0183)	(0.0178)	(0.0177)	(0.0177)	(0.0189)	(0.0181)	(0.0184)
BiProbit	-0.039***	-0.036***	-0.014***	0.003	-0.038***	-0.036***	-0.013***	0.003
	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)
Likelihood of school mobility (excluding residential mobility)								
OLS	-0.0340**	-0.0343**	-0.0323**	-0.0294**	-0.0329**	-0.0337**	-0.0317**	-0.0291**
	(0.0132)	(0.0133)	(0.0135)	(0.0132)	(0.0133)	(0.0137)	(0.0139)	(0.0138)
BiProbit	-0.033***	-0.036***	-0.014***	0.003	-0.038***	-0.036***	-0.013***	0.003
	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)
LEMO fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
School 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=2	p=2	p=2	p=2

Notes: The table presents the estimated coefficients for the window of ±4. 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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8.11 Estimation without LEMO Fixed Effects

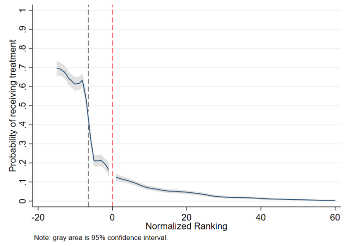
Table 11: IV Estimates with and without LEMO Fixed Effects

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Panel A: Violence Outcomes								
IV: Likelihood of Reporting Violence	0.461***	0.360***	0.461***	0.372***	0.461***	0.360***	0.462***	0.372***
	(0.110)	(0.121)	(0.110)	(0.126)	(0.110)	(0.121)	(0.110)	(0.126)
F-stat	76.55	61.02	67.02	49.41	76.53	61.03	67.00	49.43
Anderson-Rubin Test P-values	0.0000227	0.00187	0.0000695	0.00379	0.0000227	0.00186	0.0000694	0.00378
IV: Number of Reports of Violence	1.082***	0.918**	1.157***	0.995**	1.082***	0.918**	1.157***	0.995**
	(0.370)	(0.411)	(0.380)	(0.434)	(0.370)	(0.411)	(0.380)	(0.434)
F-stat	75.6	60.56	65.27	48.58	75.60	60.58	65.27	48.60
Anderson-Rubin Test P-values	0.00297	0.0212	0.00365	0.0257	0.00298	0.0212	0.00365	0.0257
N	1764	1755	1764	1755	1764	1755	1764	1755
Panel B: Dropout and Mobility								
IV: Likelihood of dropout	-0.00590	-0.00830	0.00379	-0.000257	-0.00468	-0.00698	0.00464	0.000710
	(0.00788)	(0.00889)	(0.00754)	(0.00855)	(0.00775)	(0.00867)	(0.00759)	(0.00856)
F-stat	17.47	16.22	17.29	15.33	17.14	15.43	17.15	14.76
Anderson-Rubin Test P-values	0.455	0.342	0.614	0.976	0.549	0.418	0.538	0.934
IV: Likelihood of mobility	-0.0340**	-0.0343**	-0.0323**	-0.0294**	-0.0329**	-0.0337**	-0.0317**	-0.0291**
	(0.0132)	(0.0133)	(0.0135)	(0.0132)	(0.0133)	(0.0137)	(0.0139)	(0.0138)
F-stat	17.47	16.22	17.29	15.33	17.14	15.43	17.15	14.76
Anderson-Rubin Test P-values	0.000702	0.00119	0.00117	0.00502	0.00120	0.00169	0.00198	0.00720
LEMO fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
School 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=2	p=2	p=2	p=2

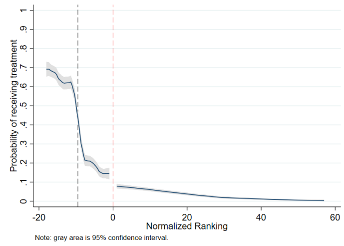
Notes: The table presents the estimated coefficients obtained after estimating our main specification (equation 4.3) for the window of ±4. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8.12 Alternative Placebo Thresholds

Figure 5: Placebo Thresholds



(a) Threshold at 15



(b) Threshold at 18

8.13 Exposure to 2018 intervention

8.13.1 2018 and 2019 Intervention

In 2018, MINEDU offered an intervention similar to the 2019 intervention. The intervention in 2018 was offered in 181 of the 221 LEMO[70] to 2605 schools in the country. In 2019, all the LEMO provided the programme and 2655 schools were targeted. 30% of the schools that received the intervention in 2019, were also exposed to at least one activity of 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 of violence. Among the schools that fulfilled one or both conditions, MINEDU prioritize the schools located near the LEMO[71] 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 data had not experienced incidents of violence. MINEDU categorized schools in two groups: nucleo schools and adjacent schools. In each LEMO, they targeted 3 nucleo schools and 9 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. They were called nucleo, as this in Spanish means core or central, and these schools represented the schools with the highest prevalence of violence located near the LEMO. After selecting 3 nucleo schools per LEMO, MINEDU selected the adjacent schools. The eligibility rule stablished that the adjacent schools would be the 3 schools located closest 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 similar distance to the nucleo school or when those close to the LEMO had few students, MINEDU prioritized the school that had a larger population of

[70]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.

[71]Approximately no more than 5 hours - by car - away from the LEMO

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.

The (apparently trivial) differences in the eligibility criteria and the scope of the intervention make crucial weighting the trade-offs between the alternative avenues to assess the impact of the interventions in 2018 and 2019. On the one hand, I can treat each intervention independently - as two different programs. On the other hand, I can treat the interventions as if they were the same intervention but implemented following a staggered roll-out. The latter option would involved using a two-way-fixed effects regressions model[72]. Yet, there are three important factors that lead me to overule this option.

First and foremost, the eligibility criteria differ across time. The criteria to select the nucleo schools and the schools treated in 2018 had similarities, yet the measures of violence that were used to select the schools were different. In 2018, MINEDU used the number of reports of violence registered in the SíSeVE platform, while, in 2019, MINEDU used a combination between the number of reports of violence and an index of perceived violence. These differences might explain that only 17% of nucleo schools treated in 2019 were also eligible and exposed to the intervention in 2018. Moreover, the eligibility criteria and process of selection of adjacent schools was substantially different to the criteria used in 2018. The other two reasons, are perhaps, less of a concern, but still are important factors to keep in mind as they would affect our interpretation of the treatment effects. In 2019, the intervention included a new topic in the curricula, as well as, group learning sessions. Finally, those treated in 2018 were not treated throughout, meaning that there is a group of schools that switch from treated to untreated[73].

Therefore, I treat 2018 and 2019 interventions independently and focus on analysing the impact of 2019 intervention. Four reasons motivate this. First, the eligibility criteria use to select the adjacent schools in 2019 provides an opportunity to evaluate the impact of the intervention in a causal way. Moreover, I focus in the adjacent schools, considering that more than 90% of nucleo school 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. Second, the likelihood of treatment in 2018 is orthogonal or independent to 2019 eligibility criteria (see section 4). Third, 4 months after the intervention was completed, I collected primary data on a sample of 2019 treated schools, allowing me to explore the implementation challenges and the school attitudes towards the intervention.

Lastly, the TA in 2019 was more homogeneous across the LEMO relative to 2018 intervention. In 2018, due to logistic constraints the intervention was not implemented in 40 LEMO. Moreover, even though the intervention was planned to be implemented between June and November, 3 months after the beginning of the academic year [74], in 21 LEMO the intervention started in October[75]. These LEMO 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.

[72]In the standard staggered adoption designs, groups adopt or receive a policy at different times and, once treated, they stay treated throughout. In this setting, researchers generally use two-way-fixed effects regressions, where the estimator of the treatment effects of the intervention is a weighted sum of the estimates for 3 groups: early adopters versus later adopters; adopters versus never adopters; and later adopters that use an already treated group as a comparison (Goodman-Bacon, 2019; De Chaisemartin and D'Haultfceuille, 2019; 2020). If this method is used, as discussed by Goodman-Bacon (2019) and De Chaisemartin and D'Haultfceuille (2019; 2020), it is essential to account for the heterogeneity in the treatment effects over time.

[73]De Chaisemartin and D'Haultfceuille (2019L; 2020) propose a method to account for situations in which treated observations switch to untreated.

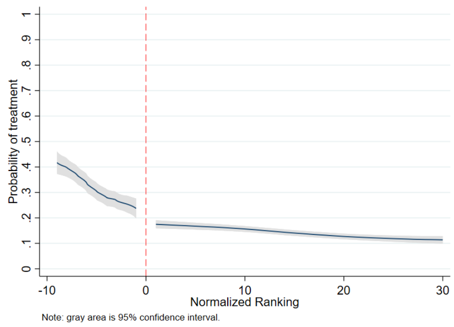
[74]The academic year in Perú starts in March and finishes by mid-December.

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

In 2019, the intervention was implemented between May and October, and the experience of 2018 allow for better planning in the implementation. Administrative data shows that two thirds of the LEMO delivered all program components and that the remaining LEMO implemented, in average, 7 out of the 9 program components.

8.13.2 Figures

Figure 6: Probability of being Treated in 2018



8.13.3 Results after dropping schools exposed to 2018 intervention

Table 12: Placebo Estimates: Reports of School Violence

	N	Control Mean	Placebo
A. Pre-treatment outcomes			
Number of Reports of Violence	1376	0.280	0.274
		(1.124)	(0.249)
First Time Reported Violence	1376	0.056	-0.024
		(0.230)	(0.064)
Proportion of students that dropout	1376	0.030	0.006
		(0.036)	(0.009)
Proportion of students that transferred to another school	1376	0.059	-0.009
		(0.053)	(0.015)
B. Covariates			
School has access to electricity, water and sanitation	1376	0.578	-0.052
		(0.494)	(0.128)
Index of School Infrastructure (material of construction)	1375	0.086	-0.139
		(1.299)	(0.327)
School Principal chosen by Meritocracy	1376	0.752	-0.135
		(0.432)	(0.118)
Proportion of teachers chosen by meritocracy	1369	0.533	-0.214**
		(0.332)	(0.095)
Proportion of parents with secondary education or more	1370	0.679	-0.042
		(0.236)	(0.047)
JEC School	1376	0.076	0.292***
		(0.264)	(0.080)

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 our main specification (equation 4.3) but using as dependent variable the outcomes and relevant covariates determined prior to the intervention in the year 2018 (results remain consistent when using 2017 as the pre-treatment year). The estimates correspond to a window of analysis of ± 4. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Violence-related outcomes

  As a robustness check, I also drop from the sample all the schools that were exposed to 2018 TA and observe that the coefficient remains the same. I only observe an increase in the standard errors that can be explained by the fact that the sample size reduces by 22% (388 schools). This finding gives more confidence that I am able to estimate the local average treatment effect of 2019 TA.

Table 13: Likelihood of reporing violence

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Sample						
IV - LATE	0.461***	0.377***	0.372***	0.462***	0.377***	0.372***
	(0.110)	(0.127)	(0.126)	(0.110)	(0.127)	(0.126)
ITT	0.209***	0.141***	0.139***	0.209***	0.141***	0.139***
	(0.0516)	(0.0479)	(0.0475)	(0.0516)	(0.0479)	(0.0475)
F-stat	67.02	49.64	49.41	67.00	49.66	49.43
Anderson-Rubin Test P-values	0.0000695	0.00354	0.00379	0.0000694	0.00353	0.00378
N	1764	1755	1755	1764	1755	1755
Panel B: Drop schools exposed to 2018 TA						
IV - LATE	0.545***	0.499***	-	0.542***	0.496***	-
	(0.131)	(0.153)	-	(0.131)	(0.153)	-
ITT	0.232***	0.178***	-	0.232***	0.177***	-
	(0.0551)	(0.0523)	-	(0.0553)	(0.0526)	-
F-stat	48.49	37.18	-	48.28	37.06	-
Anderson-Rubin Test P-values	0.0000356	0.000795	-	0.0000400	0.000878	-
N	1376	1368	-	1376	1368	-
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=2	p=2	p=2

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 ±4, 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. Columns (2) and (4) include school covariates, and columns (3) and (6) incorporate a dummy variables that takes the value of 1 if the school was exposed to activities of 2018 intervention. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Number of Reports of School Violence

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All Sample						
IV - LATE	1.157***	1.003**	0.995**	1.157***	1.003**	0.995**
	(0.380)	(0.434)	(0.434)	(0.380)	(0.434)	(0.434)
ITT	0.620***	0.373**	0.369**	0.620***	0.373**	0.369**
	(0.193)	(0.164)	(0.164)	(0.193)	(0.164)	(0.164)
F-stat	65.27	48.80	48.58	65.27	48.82	48.60
Anderson-Rubin Test P-values	0.00365	0.0243	0.0257	0.00365	0.0243	0.0257
N	1764	1755	1755	1764	1755	1755
Panel B: Drop schools exposed to 2018 TA						
IV: Number of Reports of Violence	1.193***	1.100**	-	1.195***	1.103**	-
	(0.417)	(0.466)	-	(0.417)	(0.466)	-
ITT: Number of Reports of Violence	0.572***	0.391**	-	0.573***	0.394**	-
	(0.205)	(0.173)	-	(0.206)	(0.174)	-
F-stat	47.30	36.58	-	47.10	36.47	-
Anderson-Rubin Test P-values	0.00744	0.0249	-	0.00737	0.0247	-
N	1376	1368	-	1376	1368	-
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=2	p=2	p=2

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 ±4, 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. Columns (2) and (4) include school covariates, and columns (3) and (6) incorporate a dummy variables that takes the value of 1 if the school was exposed to activities of 2018 intervention. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As an additional check, I estimate the Fuzzy RDD restricting my sample to the 40 LEMO that were only exposed to the 2019 intervention. Even though these results should be interpreted carefully as the sample of analysis reduces substantially, it is reassuring to find that the estimated coefficients remain statistically significant, the sign and direction of the estimates remain the same and the magnitude increases slightly.

Table 15: Number of Reports of School Violence - Sample of 40 LEMO

	(1)	*(2)*	*(1)*	*(2)*
IV - LATE	2.707***	2.386**	2.701***	2.390**
	(0.919)	(1.189)	(0.919)	(1.188)
ITT	1.128***	0.747**	1.124***	0.748**
	(0.383)	(0.362)	(0.383)	(0.363)
F-stat	11.52	7.674	11.51	7.624
Anderson-Rubin Test P-values	0.00539	0.0458	0.00554	0.0460
N	316	315	316	315
LEMO fixed effects	Yes	Yes	Yes	Yes
School Covariates	No	Yes	No	Yes
Polynomial	p=1	p=1	p=2	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3) in the window of ±4 for the 40 LEMO that were only exposed to the intervention in TA. 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.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Education related outcomes

Table 16: Dropout and Mobility

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Panel A: Dropout								
IV: Student Dropout	0.00217	-0.00404	-0.00419	-0.00997	0.00328	-0.00302	-0.00209	-0.00757
	(0.00726)	(0.00779)	(0.0114)	(0.0133)	(0.00733)	(0.00774)	(0.0109)	(0.0125)
ITT: Student Dropout	0.000838	-0.00151	-0.00155	-0.00332	0.00127	-0.00112	-0.000783	-0.00253
	(0.00281)	(0.00289)	(0.00403)	(0.00400)	(0.00283)	(0.00287)	(0.00400)	(0.00388)
F-stat	19.27	19.01	15.26	13.56	18.95	18.36	16.04	14.17
Anderson-Rubin Test	0.766	0.603	0.701	0.408	0.654	0.696	0.845	0.516
N	499148	495884	288293	285895	499148	495884	288293	285895
Panel B: Mobilitty								
IV: Student mobility	-0.0364**	-0.0278*	-0.0219	-0.0152	-0.0351**	-0.0270	-0.0201	-0.0137
	(0.0169)	(0.0161)	15.26	(0.0184)	(0.0170)	(0.0164)	(0.0173)	(0.0182)
ITT: Student mobility	-0.0141***	-0.0104**	-0.00813	-0.00504	-0.0136***	-0.0100*	-0.00750	0.00457
	(0.00503)	(0.00498)	(0.00570)	(0.00563)	(0.00513)	(0.00509)	(0.00577)	(0.00566)
F-stat	19.27	19.01	15.26	13.56	18.95	18.36	16.04	14.17
Anderson-Rubin Test	0.00562	0.0381	0.156	0.371	0.00865	0.0501	0.194	0.421
N	499148	495884	288293	285895	499148	495884	288293	285895
Panel C: Mobilitty (excluding residential mobility)								
IV: Student mobility	-0.0318**	-0.0239**	-0.0157	-0.00981	-0.0307**	-0.0232*	-0.0148	-0.00912
	(0.0128)	(0.0118)	(0.0134)	(0.0140)	(0.0131)	(0.0122)	(0.0134)	(0.0141)
ITT: Student mobility	-0.0123***	-0.00893**	-0.00582	-0.00326	-0.0119***	-0.00862**	-0.00552	-0.00305
	(0.00367)	(0.00361)	(0.00440)	(0.00438)	(0.00379)	(0.00372)	(0.00449)	(0.00445)
F-stat	19.27	19.01	15.26	13.56	18.95	18.36	16.04	14.17
Anderson-Rubin Test	0.000979	0.0142	0.187	0.457	0.00191	0.0214	0.220	0.494
N	499148	495884	288293	285895	499148	495884	288293	285895
LEMO fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Exposure to TA 2018 Covariate	No	Yes	No	No	No	Yes	No	No
Drop schools exposed to TA 2018	No	No	Yes	Yes	No	No	Yes	Yes
Polynomial	p=1	p=1	p=1	p=1	p=2	p=2	p=1	p=2

Notes: The table presents the estimated coefficients and standard errors obtained after estimating our main specification (equation 4.3) in the window of ±4. In each panel, the first row shows the IV-LATE coefficient estimates, while the second row shows the ITT estimates. Columns (1), (2), (5) and (6) include the whole sample, while columns (3), (4), (7) and (8) exclude the schools that were exposed at least to one activity of 2018 intervention. Robust standard errors are reported in parenthesis and are clustered at the LEMO level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8.13.4 Characteristics of schools exposed to 2018 intervention

Schools below and above the threshold that were exposed at least one activity from 2018 intervention are very similar.

Table 17: Normalized Differences between schools exposed to treatment in 2018 below and above the Threshold

	Below Threshold Exposed to Treatment in 2018	Above Threshold Exposed to Treatment in 2018	Normalized Difference
A. Pre-treatment outcomes			
Number of Reports of Violence	0.511	0.503	0.006
	(1.398)	(1.408)	
First Time Reported Violence	0.049	0.085	0.142
	(0.217)	(0.280)	
Proportion of students that dropout	0.017	0.016	0.046
	(0.026)	(0.025)	
Proportion of students that transferred to another school	0.051	0.050	0.009
	(0.039)	(0.039)	
B. Covariates			
School has access to electricity, water and sanitation	0.749	0.745	0.008
	(0.435)	(0.437)	
Index of School Infrastructure (material of construction)	0.455	0.518	0.058
	(1.067)	(1.105)	
School Principal chosen by Meritocracy	0.713	0.661	0.113
	(0.453)	(0.475)	
Proportion of teachers chosen by meritocracy	0.541	0.582	0.141
	(0.287)	(0.284)	
Proportion of parents with secondary education or more	0.599	0.593	0.027
	(0.220)	(0.209)	
JEC School	0.215	0.194	0.053
	(0.412)	(0.397)	
C. Variables related to the running variable			
Distance to Nucleo Schools	3.798	4.510	0.155
	(4.573)	(4.593)	
Total number enrolled students	596.161	481.103	0.199
	(649.826)	(496.842)	
N	*223*	*165*	

Notes: The table shows the normalized differences between schools located below the threshold that were exposed to treatment in 2018 and 2019 and schools located above the threshold that were exposed to treatment only in 2018. Normalized Differences larger than 0.25 indicate that the average covariate values are different between the two groups.