# Violence and human capital investments

Foureaux Koppensteiner , Martin; Menezes, Livia

*DOI:*
10.1086/711001

# Violence and Human Capital Investments*

Martin Foureaux Koppensteiner† Lívia Menezes‡

July 16, 2020

**Abstract**

In this paper, we combine extremely granular information on the location and timing of homicides with a number of large administrative educational datasets from Brazil, to estimate the effect of exposure to homicides around schools, students' residence, and on their way to school. We show that violence has a detrimental effect on school attendance, on standardised test scores and increases dropout rates of students substantially. We use exceptionally rich information from student- and parent-background questionnaires to investigate the effect of violence on the aspirations and attitudes towards education. We find that boys systematically report lower educational aspiration towards education.

*We would like to thank Esteban Aucejo, Fernanda Brollo, Paolo Buonanno, Antonio Ciccone, David Figlio, Jeffrey Grogger, Randi Hjalmarsson, Mark Hoekstra, Michael Lovenheim, Olivier Marie, Philip Oreopoulos, James Rockey, and Ian Walker and seminar participants in Birmingham, Brazilian Econometric Society (SBE), CEP (LSE), Essex, Lancaster, Leicester, Reading, Surrey, Sheffield, St. Andrews, the 2017 IZA Workshop on The Economics of Education, the 2018 Workshop on Human Capital at the Bank of Italy, the 2018 RES Annual Conference, the 2018 EEA Annual Congress, the 2018 ESPE annual conference, the 2018 York Workshop of Labour and Family Economics, the 2018 AEDE meeting, the 2018 EALE Annual Conference, the 2018 Briq Workshop Bonn, the 2019 Southampton Workshop of Applied Microeconomics, and the 2019 North American Winter Meeting of the Econometric Society for their very useful comments and suggestions. We also thank Onigroup for their support with geocoding the data. Our particular thanks goes to Kerstin Pittl for her support with the geocoding processes and the creation of school corridors. We also thank Tomislav Bacinger for his support with the creation of interactive homicide maps. We also express our gratitude to the Brazilian Ministry of Health, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) and the State Secretariat for Education in São Paulo for the data used in this project. Foureaux Koppensteiner gratefully acknowledges financial support received from British Academy grant (SG150627). This research used the ALICE High Performance Computing Facility at the University of Leicester. The usual disclaimer applies.

†University of Surrey, Guildford, Surrey, GU2 7XH, UK and IZA, Bonn, Germany. Email: m.koppensteiner@surrey.ac.uk. Phone: +44 14836 83995

‡University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. Email: l.menezes@bham.ac.uk.

# 1 Introduction

According to statistics from the World Bank, Brazil has one of the highest homicide rates in the world. In 2016, the intentional homicide rate in Brazil was more than 29 per 100,000 people, which is approximately 6 times the US rate and 29 times the UK rate and according to national security statistics, in 2016, 61,283 homicides were registered in the country.[1] The Brazilian Institute of Applied Economic Research (Ipea) estimated that the cost of violence corresponds to more than 5% of the country's gross domestic product (GDP), not including many intangible costs that are difficult to quantify (Cerqueira et al. (2007)). The pain, suffering, and trauma caused by direct victimisation and exposure to violence in the local neighbourhood may negatively affect a variety of societal outcomes, among those educational production, but little is known about the causal effect of exposure to day-to-day violence on educational outcomes. In this paper, we estimate the effect that exposure to violence has on the human capital accumulation of students in Brazil, using a unique novel dataset containing georeferenced information on all homicides occurring in the public way and combining this with very detailed information on a wide range of educational outcomes, including educational aspiration of students.

Violence may affect school supply and the behaviour of students, parents, teachers, and principals. Several qualitative studies by psychologists, psychiatrists, and sociologists have found a range of adverse consequences in the behaviour of children after exposure to community violence: depression, anxiety, hyper-vigilance, avoidance, aggressive behaviour, delinquency, and deterioration of cognitive performance (Cooley-Quille et al. (1995), GormanSmith and Tolan (1998), Fowler et al. (2009), Farrell and Bruce (2010), Sharkey et al. (2014)). Community violence can also affect attendance at school. When a crime occurs in the neighbourhood or in the proximity of the schools, parents may feel uneasy about sending their children to school. According to the 2012 edition of the Brazilian National Survey of School Health[2], almost 9% of the 9th grade students that answered the survey declared they had stopped going to school at least once in the 30 days preceding the survey due to not feeling safe on the way from their residence to school. Low attendance may then damage the learning outcomes of students by missing curriculum content and regular contact with teachers and classmates. Exposure to homicides may also reveal information to students and parents about likely victimisation and affect the expected returns on education and hence the optimal schooling

[1] www.forumseguranca.org.br.

[2] *Pesquisa Nacional de Saúde do Escolar*, in Portuguese. Available from www.ibge.gov.br.

decision.

Because of the potential for such negative externalities, the cost of violence may go well beyond the cost of direct victimisation. Poor neighbourhoods with lower socio-economic status often register higher rates of violence, and if this also has a negative effect on human capital accumulation, this could be a relevant channel leading to the perpetuation of poverty.

This paper estimates the causal effect of exposure to violence on schooling performance, using a unique set of Brazilian microdata. Different from previous studies, which have mostly focused on variation from secular changes in violence, examples include the Mexican war on crime, conflict between rivalling gangs or armed conflict, we focus on violence related to crime more generally. The variation from this day-to-day violence stands in sharp contrast – both with respect to frequency and the spatial distribution – to the variation from secular changes to levels of violence, which often comes along with changes in possible confounders. We have information on the exact date and precise location of each homicide in Brazil occurring between 2007 and 2013, and combine this with the location of the schools and residences of students. We exploit the variation of homicides across space and over time to estimate the effect of exposure to homicides on a number of educational outcomes including test scores, repetition, dropout rates, school transition, and attendance, while controlling for school and time fixed effects, and in the most satiated specifications, school-specific time trends.

There are few studies estimating the relationship between exposure to day-to-day violence and school performance (e.g. Grogger (1997), and Aizer (2008)), which given the cross-sectional nature of their data generally cannot deal with the endogeneity problem arising from the fact that violence might be correlated with other sources of socio-economic disadvantages and school outcomes. McMillen et al. (2019) focus on a different angle and estimate the effect Chicago's Safe Passage programme of placing guards along specified school paths on crime and school absenteeism. There is also a related literature focusing on violent conflict, making use of variation in conflict across space and over time to estimate the effect of conflict exposure on educational outcomes. Brück et al. (2019) study the effect of the Israeli-Palestinian conflict on various education outcomes for Palestinian high school students by exploiting within-school variation in the number of conflict-related Palestinian fatalities during the academic year. Brown and Velásquez (2017) use the secular change in violence induced by the Mexican war on drugs to estimate the effect on human capital accumulation using municipality-level changes in drug-related violence. They argue that the negative effects of violence on education are driven by the economic consequences and financial hardship to households in relation to the violent conflict. Monteiro and Rocha (2017) estimate the effect of gunfights between drug gangs in Rio de Janeiro's favelas (slums) on student achievements using panel data for the city of Rio de Janeiro, examining the effect of conflicts in favelas on students who study in schools located in favelas and in their close proximity.

Violence in Brazil is nevertheless a more widespread phenomenon that differs from armed conflict between drug gangs occurring in favelas, in terms of the intensity and the concentration of occurrence, both over time and across space. The measure of violence we use, homicides, captures the widespread nature of violence in Brazil and allows us to estimate the effect of day-to-day violence on student achievements in a much more general context, likely to be much more representative of the violence Brazilians face on a daily basis. The extremely granular information on the location of homicides allows us to investigate the exposure in the very close proximity to schools and the residences of students.[3] This makes the results presented in the paper relevant for the understanding of the externalities of day-to-day violence present in Brazil and in many other countries.

We focus our analysis on the city of São Paulo, which is the largest city in the Americas with a population of 12 million people. São Paulo provides an ideal setting for our study, because of the extremely detailed schooling outcomes we have available for São Paulo, and because of the sheer size of the data. The city provides also an interesting case study for understanding day-to-day violence in countries with more moderate crime levels, as it ranks close to the US in terms of the homicide rate.[4]

There are three main contributions of the paper. First, we provide credible causal estimates of the effects of day-to-day violence on schooling outcomes. For that purpose, we combine extraordinarily rich set of microdata on student outcomes with a measure of violence that is consistent across space and time: homicides. These allows us to focus on variation in day-to-day violence over time and across space, that is comparable and minimises measurement error. For these homicides, we have extremely granular geocoded address information, which we match with information on the addresses of the schools and residences of students attending these schools. This allows us to investigate the effect of exposure to violence around the schools and residences of students. We find that violence around the schools leads to a substantial deterioration in the educational performance

[3] The context also likely affects the reporting of violence, which is why we focus our analysis on violence measures based on official death records, minimising the risk for selective reporting.

[4] The homicide rate in São Paulo has dropped dramatically over the past 20 years, from highs around 53 to 6 per 100,000 population in 2018, roughly equivalent to the level of the homicide rate in the US.

of schoolchildren. We find that one additional homicide in a 25 m radius around schools reduces test scores in math and language by about 5% of a standard deviation in test scores. Furthermore, we find that homicides increase dropout rates and have a negative effect on attendance. We also find that the effects are particularly pronounced among students from relatively poorer families, indicating that income may work as a buffer against the negative effects of crime. We also show that the estimated effects are particularly strong for boys, both for test scores and attendance.

Second, we systematically investigate exposure to violence on the way to school for each student. We use Google APIs and design an algorithm to build corridors along the path from the residences to schools and investigate the effect of exposure to homicides in these corridors. As corridors vary in length for students attending the same school, different corridors mechanically may have different propensities to experience homicides. To deal with this, we provide within-corridor estimates and find a substantial and economically meaningful increase in dropout rates as result of exposure to violence on the school path. An additional homicide leads to an increase in the probability of a student dropping out of school of 3%, an increase of about 20% compared to the baseline.

Third, we make use of the extremely rich information we have available on students, teachers and parents to investigate the underlying transmission channels. We show that different from the armed conflict settings in Monteiro and Rocha (2017) and Brück et al. (2019), day-to-day violence does not affect the supply of schooling, either through teacher absenteeism, teacher or head teacher turnover. We find that the effects are not driven by short-lived effects before the test date or by reductions in school attendance only. We use the extremely rich information we have available on student and parent reported educational aspirations and attitudes and find that boys' aspiration suffers as a consequence to homicide exposure. This is consistent with a transmission channel that operates through a differential effect on the incentives to invest in human capital for boys, who make up the vast majority of victims in homicides in Brazil. This paper hence also contributes to the literature that investigates how life expectancy affects human capital investments (Jayachandran and Lleras-Muney (2009) and Gerardino (2014)).

## 2 Institutional Background

The Brazilian educational system is predominantly regulated by the federal government, which is also responsible for distributing resources to states and municipalities. These secondary layers of government not only manage the funds received but are also allowed to implement state- or municipality-specific programmes and policies. The educational system is composed by two main levels: 1) *Educação Fundamental* (basic education), which comprises *Educação Infantil* (nursery), *Ensino Fundamental* (primary school), and *Ensino Médio* (secondary education) and 2) *Educação Superior* (higher education).

Public primary education is offered at no cost for all, irrespective of age, and it is mandatory for children between 6 and 14 years of age. It lasts nine years,[5] and it is divided into two stages: the first cycle, which comprises 1st to 5th grade, and the second cycle, which includes 6th to 9th grade. Public secondary school is also offered at no cost and lasts 3 years. It is not compulsory, but recent regulation pushes towards gradually making secondary education compulsory as well. To be able to enrol in secondary school, students must conclude primary school.

A school year contains at least 800 hours spread over at least 200 school days. The precise starting and ending days of the school year vary across schools and over the years. Figure E1 in the appendix exemplifies the school calendar in São Paulo for 2010. Every year, the São Paulo State Secretariat of Education formally announces, by releasing a document called *Resolução*, the desirable starting day of the school year. In general, the first semester finishes on the last working day of June. The second semester starts on the first working day of August and finishes on the last working day before Christmas. Each semester is composed of two bimesters, with roughly 50 days each. The precise ending dates of each bimester are school specific. This setup leads to semesters that are defined state-wide, and bimesters that are school specific. Students may be retained in a grade at the end of the year if they do not achieve adequate school performance and/or they do not meet the minimum level of attendance required by law, which is at least 75% of the school days in primary schools and 85% in secondary schools.

Considering the nature of funding and administration of schools, they can be classified into four types: federal, state, municipal, and private schools. The first three are essentially public schools, maintained by the respective administrative units. In general, private schools are of better quality; however, only a relatively small share of the population can afford the substantial school fees charged by these schools. At least 87% of the students go to public schools in Brazil. In São Paulo, this number is slightly smaller at 80%. Schools may offer all or only specific levels of basic education, and there are schools that offer only primary education, some only secondary education, and some offer both.

[5]Previously, primary school began at age 7 and lasted eight years. In 2006, the government passed a law that expanded primary school from eight to nine years and mandatory enrolment at 6 years old. States and municipalities had until 2010 to implement the new law.

Public school students are not bound to a specific school. They are able to enrol in any school with vacancies. In most cases, students attend schools located within walking distance of their residences. When this is not possible, they may qualify for school transport.

# 3 Data

We build a novel dataset by combining administrative data from three institutions: the Brazilian Ministry of Health, Brazilian Ministry of Education, and São Paulo State Secretariat of Education, and link these datasets using school, class, and individual identifiers and geographic information from the addresses.

## 3.1 Educational data

We focus the analysis on the city of São Paulo over the period from 2007 to 2013. For the educational outcomes, we combine three different datasets, the Brazilian school census collected by INEP[6] on behalf of the Brazilian Ministry of Education, data from SARESP[7] containing test score data and student and parental background information, and individual attendance records of students in all state schools, which are collected by the São Paulo State Secretariat of Education. Below we will discuss in turns the different datasets and outcome variables we use in this paper.

The annual Brazilian School Census contains rich information on the universe of students in primary and secondary school, based on individual records schools are obliged to collect on their students. Unique student identifiers allow us to follow students over time and across schools, which enables us to construct some of the outcomes we use in the analysis: grade repetition, dropout and transition from primary to secondary school. Characteristics of students and teachers include date of birth, sex, race, the grade students attend, and teacher educational background (among others). For all schools we have access to their precise addresses, which we geocode using Google Maps API.

For all students attending municipal schools, we have unique access to information on the address of the residence of the students. Unfortunately, this information is not available for students in state or private schools, as it is not collected by the State Secretariat of Education for the school census. This means we are bound to restrict any analysis involving exposure around the residence

[6] *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*, namely the National Institute for Educational Studies and Research "Anísio Teixeira".

[7] *Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo*, namely the education evaluation system of the state of São Paulo.

to students attending municipal schools. For consistency, we do not consider nursery schools[8] or any kind of special education, which is offered to students with special needs.

The majority of observations cover students in primary school (84%).[9] Measures of school efficiency, such as repetition and dropout rates, reveal substantial problems in the Brazilian educational system. More than 6% of schoolchildren repeat any given grade, and almost 10% drop out of a given grade.[10] In terms of transition from primary to secondary education, around 75% of students carry on beyond compulsory education and enrol in secondary school.

We combine the records from the Brazilian School Census with data from SARESP, which contains both, standardised test scores in math and Portuguese, and student and parent background questionnaires, using unique student identifiers. The exam is carried out every year and evaluates the performance of students in Portuguese and math in the 5th, 7th, and 9th grades of primary and in the 3rd grade of secondary school.[11] To be able to compare the results to national standardised exams, we focus on test scores for the 5th and 9th grades of primary school and the 3rd grade of secondary school. These coincide with the end of each of the educational cycles and allow us to investigate the effects for different grades and ages, including both primary and secondary school students. Test scores are normalised to a mean of 250 and a standard deviation of 50 allowing us to compare the effects across the different grades.

SARESP collects additional information through very detailed student and parent background questionnaires. These are completed after the exam by the students and are taken home and completed by the parents or legal guardian. For this paper, we are particularly interested in the children's self-assessment of their performance, their educational aspirations and involvement in school, the parental assessment of their children's performance and involvement, the parental self-assessment of their involvement, and the assessment of parental involvement by their children. For instance, for the self-assessment of their involvement in school, children are asked whether they do their homework, whether they plan to go to university, or whether they perceive themselves as good students. Parents answer questions on their perception of their children's engagement in school and on their on their involvement with their children's education.

[8]Pre-primary education has gone through a period of very rapid expansion over the last years and comprises a number of different levels across ages, which makes it difficult to come up with a consistent definition of pre-school type.

[9]This number reflects both, the longer period of primary education, 9 years versus 3 years of secondary education, and the non-compulsory nature of secondary education.

[10]We define repetition as a student being enrolled into the same grade in the following year. The variable dropout includes the temporary dropout rate, where students leave school for one or more years but enrol at school again at a later point. The variable also includes students who do not enrol in secondary school after leaving the school system after primary school.

[11]SARESP is mandatory for schools in the state system, but municipal and private schools can opt-in and participate in the tests.

We also have access to unique attendance records for students in state schools from the *Boletim Escolar*[12][13] The dataset provides the individual attendance records of all students at state schools in São Paulo at a bimonthly frequency reported as the number of school days missed. For the analysis we aggregate the bimonthly attendance records at the semester level.[14]

### 3.2 Violence data

To create our measure of localised violence, we use microdata of official death records published by the Brazilian Ministry of Health. This dataset comes from the Mortality Information System[15], which compiles information from the universe of death certificates on all natural and non-natural deaths in Brazil. We use information from the ICD-10 coding of cause of non-natural deaths to identify victims of intentional homicides. In addition to cause of death, the death certificates contain characteristics of the deceased, such as date of birth, sex, race, occupation, and the location of occurrence of the homicide. We exclude deaths to use of lethal force by law enforcement officers.[16]

We have information on the precise location available for 2,470 homicides that occur in the public way over the period of 2007 to 2013 in the city of São Paulo. These are homicides that occur openly and visibly in the public space, and hence exclude homicides occurring in private space, for example in residences.[17] We believe these homicides are particularly salient for our analysis for two reasons. First, these homicides garner considerable attention and are clearly visible to the population. Second, these homicides form a more homogeneous group (and largely exclude domestic homicides). We geocode homicide addresses using the Google Maps API and restrict our data to homicides geocoded at the street level, which correspond to 95% of all homicides in the public way.

Table E1 in the appendix displays summary statistics of the victims of homicides for which the death occurs in the public way, and the description of the characteristics of homicides. Ap-

[12] Namely the School Bulletin.

[13] These attendance records are of excellent quality and differ from attendance records of *Sistema Presença* collected for checking conditionality for the conditional cash transfer programme *Bolsa Família*. First, the attendance records from *Boletim Escolar* are available for all students, not only for recipients of *Bolsa Família*. Second, because of the purpose to inform parents about the school attendance of their children, the school administrators or teachers have no incentive to manipulate the attendance record in contrast to the *Sistema Presença* records (Brollo et al. (2019)).

[14] This is because the start and end dates of the bimesters vary by school and these dates are not centrally available. Figure E1 in the appendix presents the timing of the school calendar in São Paulo.

[15] Sistema de Informação sobre Mortalidade.

[16] These deaths make up only a very small fraction of intentional homicides in the city of São Paulo, and indeed there was not a single occurrence in the public way for the time period used.

[17] Access to address information on homicides occurring in private space was denied due to the risk of identifiability of the victims. In Figure E2, we present a comparison between homicides in the public way and the remaining homicide cases, which may occur at a hospital or residence, for example.

proximately 70% of the homicides are a result of assault by gun discharge, and about 10% each by assault using a sharp or blunt object. The majority of victims are in the age group between 19 and 50 years old, but a substantial number (8.4%) of relatively young victims of homicide are between the ages of 11 and 18. The vast majority of victims are male, and individuals from a lower socio-economic background are over-represented as victims of homicides, as indicated by the very low levels of completed education. Figure E3 in the appendix shows the distribution over time and space of the homicides in the public way in São Paulo. Darker shades of red represent areas more affected by homicides. In the paper, we use the variation of homicides over time and space depicted in the maps, allowing us to disentangle the effect of violence from other correlates of socio-economic variables and thus establish causality between violence and education, as described in the next section.[18]

## 4 Identification Strategy

Disentangling the effect of violence on education from confounding factors is not straightforward. For example, when investigating the effect of neighbourhood violence on educational outcomes, relatively poorer neighbourhoods may register higher homicide rates, and students from disadvantaged backgrounds may also be more likely to attain unsatisfactory results at school, leading to a positive relationship in these variables even in the absence of any causal effect of violence on education. Hence, it is necessary to deal with confounding factors that may lead to a positive association between levels of violence and poor educational performance are responsible for the estimated results. In addition, when using homicide rates for geographic areas, for example the municipality or neighbourhood, it is possible that a negative trend in school quality in the area, for example through lower investment in local schools, leads to an increase in crime in the area, for instance through an increase in the number of students dropping-out of school and joining gangs leading to reverse causality. It may also be possible that a deterioration of student quality in a neighbourhood, leads to an increase in drug consumption and violence, and hence reverse causality when estimating the effect of neighbourhood violence on schooling outcomes.

In this paper, we use variation in homicide exposure at a much more disaggregated level, rather than homicides aggregated at the neighbourhood or municipality level that may be subject to the above confounders. In detail, we use localised variation in homicides over time, making use of the

[18] A dynamic homicide map locating all homicides in the public way for all of Brazil and the period from 2003 to 2016 is available on www.cost-of-crime.com/homicide-map.

precise location of these homicides at the exact street address: for a large part of the analysis we focus on extremely granular exposure points, namely homicides occurring in a 25 m radius around schools.

Using the variation in homicides in the very vicinity of schools, we estimate the effect of exposure to violence on educational outcomes using the following estimation equation:

$$y_{ist} = \beta_0 + \beta_1 homicides_{st} + X_{it}\beta_2 + Z_{st}\beta_3 + d_s + d_t + d_{st} + u_{ist}, \tag{1}$$

where $y_{ist}$ is a range of different measures for the educational outcomes of $student_i$ in $school_s$, $homicides_{st}$ is the number of homicides that lie in the close periphery of schools, $X_{it}$ is a vector of individual characteristics, $Z_{st}$ are school and classroom time-varying characteristics, $d_s$ and $d_t$ are school and time fixed effects, respectively, and $u_{ist}$ is an error term. We also estimate models including school-specific time trends, denoted as $d_{st}$, probing the robustness of the estimates.[19]

For identification, we assume that, conditional on time and school fixed effects, the variation in the number of homicides in the very small geographic area around schools in a given period is random. School fixed-effects effectively control for any unobserved time-invariant school characteristics and the general composition of students in schools based on the school catchment area. In addition, school specific time-trends control for exogenous factors affecting changes in schooling outcomes and deal with any school-specific time-varying unobservable characteristics. We also include a very rich set of individual, teacher, classroom, and school characteristics to reduce error variance. The inclusion of these controls should not affect the estimates in a meaningful way, given our identification strategy, as we are in practice holding the socio-economic composition of students (and school inputs) constant. Although our identification strategy does not rely on baseline characteristics being balanced across exposed and non-exposed schools, given the very localised measure of exposure to homicides, we can still directly test for this. For this purpose, we define schools that, over the period of interest, are exposed as 'ever exposed' schools and all others as 'never exposed' schools. Results in Table E2 show that student characteristics are balanced across a very large number of socio-economic background variables. We find that school characteristics are very similar, and for a very large number of variables, only three differences are statistically significant, in line with expectations.[20]

[19]Because test score data is not available annually, but only at the end of 5th and 9th grade for primary and end of secondary school, we cannot estimate models including prior achievement or include student fixed effects.

[20]Alternatively, we also regressed 'treatment status' on the full set of individual and school characteristics using a school panel setup, and we cannot reject the null hypothesis of no joint significance in an F-test.

We define *exposure to violence* as a homicide occurring within a 25 m radius around the geo-referenced school addresses.[21] There are four main advantages of using this very granular measure of exposure to violence: first, conditional on school fixed effects and time trends, and, in the most satiated specifications, school-specific time trends, exposure of students to violence can credibly assumed to be conditionally random. Second, the count of homicides in the very close vicinity of schools or the home address of students constitutes a very salient measure of exposure, because it is likely that students attending an affected school or living at an address either directly or indirectly observe the violent act. On aggregate, this leads to a more well defined measure of exposure for students at a given school, avoiding to group together the exposure to homicides very close by and further away. This differs from measures based on homicide rates at the neighbourhood or municipality level for which it is less likely that students were directly exposed, making the interpretation of the coefficients based on aggregate violence measures more cumbersome. Third, the exposure to violence based on a small radius around schools minimizes issues with using the same homicide in the exposure measure to more than one school, whereas a larger radius may lead to a measure of homicide exposure that overlaps for different schools.[22] Fourth, the very granular measure also limits double exposure over time. For the 25 m radius measure, only about 6% of schools are exposed more than once in the seven year period.

Using our granular measure of exposure of 25 m comes at the cost of using only a fraction of the 2,470 homicides over the period. For this reason, we also vary the radius and create exposure measures for 100 m and 500 m around schools and residences as robustness. We expect that the effect of homicides further away from schools and the residences have less of an impact. To test this more directly, we also create exposure measures defined as rings when expanding the radius, by excluding homicides in the prior, smaller radius. For these rings, the dilution of the effect should be more direct and we expect the effects to diminish at a faster rate. For the smallest radius of 25 m we have 58 homicides (for 54 schools exposed), rising quickly to 268 and 1,935 homicides for a radius of 100 m and 500 m, respectively. Naturally, due to the much larger number of addresses, for the exposure measure around residences, we make use of a much larger number of homicides, 352, 1,182 and 2,274 homicides for a radius of 25, 100 and 500 m, respectively.

We present an example of the distribution of homicides and schools in the maps in Fig-

[21] In the appendix we provide details on the geocoding processes and the construction of the different exposure variables.

[22] For our preferred radius of 25 m, we have a minimal overlap, and 97% of homicides are unique to one school. This overlap quickly increases with larger radii: for 100 m, we have 89% unique homicide-school combinations, and the number drops to less than 7% for a radius of 500 m.

ure E4 in the appendix. Each individual map shows schools and homicides in the public way in a neighbourhood in São Paulo in a semester. These maps are also available on our website https://www.cost-of-crime.com/homicide-map as fully dynamic maps displaying the variation of homicides over time and across space, and including the location of schools in Brazil.

In addition, we are interested in estimating the effect that exposure on the path from the residence to school has on educational outcomes. For this purpose, we created corridors around the shortest distance path from the home address of students to their schools using Google APIs. In effect, we built polygons of varying orthogonal width from the path and count the homicides occurring within these corridors. In line with the school and residence measure, we create corridors of 50 m width centred on the path, matching the 25 m radius of the school and residence measure. We also estimate corridors for 100 m widths to test the sensitivity of the estimates. An attractive feature of these corridors is, that because of their near complete spatial coverage, we make use of a very large fraction of the entire number of homicides; we use 2,357 homicides for a 50 m width, and 2,365 for a 100 m width. As these corridors are different for students attending the same school, but living at different addresses, we estimate a variant of Equation 1:

$$y_{ict} = \beta_0 + \beta_1 homicides_{ct} + X_{it}\beta_2 + Z_{st}\beta_3 + d_c + d_t + u_{ict}, \tag{2}$$

where $y_{ict}$ denotes the educational outcomes for students in the same corridor, hence attending the same school and living at the same address/postcode.[23] $homicides_{ct}$ is the number of homicides occurring in a corridor during a school year, and $d_c$ is a corridor fixed effects. Because corridors vary in length for students attending the same school, different corridors mechanically have a different propensity to be exposed to homicides. Corridor fixed effects hold this propensity constant over time, effectively eliminating the mechanical difference for exposure. Because different corridors lead to the same schools, the model including corridor fixed effects effectively also holds school time-invariant characteristics constant. Alternatively, we also estimated models including time and school fixed effects while controlling for distance.[24] The estimates of these models are very similar to the estimates of the corridor fixed effects (results available in Table E19 in the appendix). We calculated three alternative corridors: walking, driving, and public transport. Rather than seeing these strictly as the walking, driving, and public transport path, we consider these simply

[23] There are multiple observations from following the same corridor over time and from the fact that multiple students live at the same address and attend the same school.

[24] We use the natural log of the calculated path distance from the Google Directions API. We restrict the maximum corridor length to 3,000 m to limit the number of API calls and to exclude cases of mistaken address information from geocoding.

as alternative corridors useful to determine the sensitivity of a particular path. Figure E5 in the appendix shows a fictional example for a walking path corridor including different widths and exposures to homicides on the path. We present and discuss the results from this exercise in the appendix.

# 5 Main results

In this section, we present the results of the effect of exposure to violence around schools using the granular 25 m radius exposure measure of homicides. We start with estimating the effect of homicides on measures of academic achievement in Subsection 5.1. We then investigate heterogeneous effects in Subsection 5.1.2. Finally, we estimate the effect of homicide exposure on a number of additional outcomes, including school attendance, self-reported measures of aspiration, attitudes, and perception of students and their parents, and measures of student progress through the education system.

## 5.1 Effect of homicides on academic achievement

First, we estimate the effect of exposure to violence on academic achievement using the standardised test scores in math and Portuguese from SARESP. As the explanatory variable *Homicides*, we use the count of homicides in a 25 m radius around the school. We present robust standard errors clustered at the school level in parentheses. To account for possible spatial dependence among schools and for serial correlation, we also compute Conley standard errors[25] (Conley (1999)), presented in brackets. Table 1 presents the regression results of the effect of violence on math and language test scores for students in the 5th and 9th grades of primary school and the 3rd grade of secondary school.

In the first column, we estimate the effect of homicides on standardised math test scores, including school and time (year) fixed effects without further individual or school controls. In the second column, we include the rich set of student, teacher, classroom, and school characteristics as controls.[26] In the third column, in addition to the full set of controls, we include as a control the interaction between school and time, allowing for school-specific time trends.

Across specifications, we find a negative effect of homicides on math test scores. Adding the

[25] We compute Conley standard errors using a 25 m cut-off distance in accordance with the definition of measure for exposure. Results remain the same for 50 m or 100 m.

[26] Please see table notes for a detailed description of the full set of controls.

full set of controls does not significantly change the estimate, lending further credibility to the identification strategy. The inclusion of the controls nevertheless reduces noise and hence improves precision of the estimate. Using our preferred specification in column (2) including the full set of controls, we find that an additional homicide in the surroundings of schools during the year decreases math test scores by about 2.3 points, an effect equal to roughly 5% of a standard deviation of test scores.

The inclusion of school-specific time trends increases the results slightly, to an effect size equal to 6.6% of a standard deviation, significant at the 5% level.[27] Applying Conley standard errors to address potential spatial and serial correlation reduces the standard errors and improves precision further, suggesting that spatial and/or serial correlation of homicides is not relevant in our context.

In columns (4) to (6), we repeat the exercise for Portuguese language scores. Across specifications, we find that exposure to homicides around schools has a negative effect on test scores of slightly smaller but overall roughly similar magnitudes.[28] The coefficient for our preferred specification in column (5) is 2.1 percentage points, equivalent to about 4% of a standard deviation in test scores.

We use the information on the timing of homicides for a falsification exercise. Mechanically, homicides occurring after the test dates should not affect the test performance of children. To test for this, we create one year leads of our explanatory variable. A significant effect of the lead homicide measure may indicate a violation of the identification assumptions. In columns (5) and (10) of Table E5, we report the coefficients for the lead exposure variable. We find no effect of homicide leads on either math or Portuguese language test scores. The coefficients are much smaller and not statistically significant, lending extra credibility to our estimation strategy. [29]

The estimated effects on math and Portuguese test scores are sizeable and economically important. To put our estimates in context, we suggest comparing the effect of exposure to one additional homicide with the effect of educational inputs, for example teacher quality. Our estimates show that exposure to a homicide in the school vicinity has approximately the same effect as a reduction in teacher quality[30] by half a standard deviation on nationally standardised distri-

[27]When comparing the coefficients pairwise across our specifications, none of the pairwise differences are statistically significant.

[28]This is consistent with the findings of Monteiro and Rocha (2017), who found that the coefficients for language are generally smaller compared to the effects for math test scores. Our results differ though from their estimates on language test scores, as their estimates for the Portuguese language are a magnitude smaller and not significant at conventional levels.

[29]We repeated the exercise for additional outcome variables, including attendance and dropout. The coefficients on homicide leads for these outcomes are small and not statistically significant. Results are available upon request from the authors.

[30]As estimated by Rockoff (2004).

butions of achievement, demonstrating the economic relevance of the effects. With violence being a widespread phenomenon in Brazil and homicides reaching an all-time high in recent years, this suggests that exposure to violence may contribute significantly to low achievement of students, particularly in areas more prone to violence. While São Paulo offers an ideal setting to study the effect of violence on homicides because of the outstanding educational data, it is indeed the state with the lowest homicide rate in Brazil,[31] making our estimates potentially even more relevant for states with higher homicide rates and a higher propensity for exposure.

This nevertheless also raises the question regarding how the effect of exposure to a homicide varies by general crime levels. More frequent exposure to crime may either lead to a stronger or weaker reaction to homicide exposure. To test for how our effects vary by crime levels, we estimate our preferred specification from columns (2) and (5) in Table 1 separately for children in schools in *high-crime* and *low-crime* areas.[32] We present the results in Table 2. We find that the effects are much more pronounced in *low-crime* areas, both for math and Portuguese test scores. We find no effect in *high-crime* areas.[33] This is consistent with the hypothesis that the effects of violence are relatively less pronounced when violence is endemic, a result also documented in another context, when estimating birth outcomes of mothers affected by homicide exposure in Brazil (Foureaux Koppensteiner and Manacorda (2016)).

To understand whether the main effects are driven by shifts in the lower or upper part of the test score distribution, we also create indicator variables classifying students into proficiency levels. For both math and Portuguese performance, we create variables indicating whether a student's test result was *very low* (10th percentile), *low* (25th percentile), *median*, *high* (75th percentile), and *very high* (90th percentile). The variables *high* and *low* correspond to what the State Secretariat defines as the 'advanced' level and 'below the basic level' of proficiency. Table 3 presents the results. We find that exposed students are more likely to be classified as performing at *very low* and *low* levels of proficiency, both in the math and Portuguese language tests, but the coefficients are not statistically significant at conventional levels. We also find that students are less likely to be classified as performing at *high* and *very high* levels, indicating that students over the entire test score distribution are affected by homicide exposure, but again the estimates are not statistically

[31] www.forumseguranca.org.br

[32] For this purpose, we consider the homicide count in rings around schools (between 500 and 25 m radii around schools) over the entire period (2007-2013) to proxy for the homicide proneness of the school surroundings. We classify school surroundings as *low* homicide, where the homicide count is less than or equal to the median and as *high* where the homicide count is above the median.

[33] Alternatively, we use only homicides from 2007 to 2009 to define the neighbourhood crime level, and the estimates are basically unchanged. Results available from the authors.

significant. The shift in the distribution is nevertheless more pronounced for below the median in the test score distribution.

#### 5.1.1 Robustness checks

In the online Appendix B we provide a battery of robustness checks. We first show how the effects vary by allowing for different radii of homicide exposure around schools and we provide estimates based on exposure rings of differing width. Next we provide estimates for different timing of homicides from the end of the year tests. We also provide a falsification exercise using lead homicide measures. Third, we investigate how the effect varies by characteristics of the homicide victim, and lastly we show that the effects on test scores are not affected by selection into the test.

#### 5.1.2 Heterogeneous Effects

##### 5.1.2.1 Analysis by gender

In Table 4, we present the results of the effect of violence in the school surroundings on math and language standardised test scores separately for boys and girls. All specifications include time and school fixed effects and the full set of controls. We find negative effects of homicide exposure for both boys and girls, but the effects on boys are more pronounced than for girls. For each additional homicide around the school in the year, boys' math proficiency decreases by about 5.9% of a standard deviation and their Portuguese language proficiency decreases by 5.6%. The effect on girls is about half this size for math at 3.5% of a standard deviation in math, and only significant at the 10% significance level when considering Conley standard errors. Girls' language coefficient is not significant at conventional levels.

Strikingly, while we find more pronounced effects on educational outcomes for boys compared to girls, we find that parents evaluate the safety of their children at schools differently. We present these estimates in Table E10 in the appendix. Asked about whether parents think their *child is safe at school* or *feels safe at school* and about their rating of the security at school, parents perception of the safety of their children is reduced throughout all of these categories for boys and girls. The effects are nevertheless much more pronounced for girls, suggesting that the subjective evaluation of school safety by parents suffers more for girls than boys. This suggests that the stark difference we document in the effects on math and Portuguese test scores are not driven by the relative shift in the perception of safety (by parents).

These stark differences by gender may be an indication of fundamentally different transmission mechanism, related to changes to the incentives to invest in human capital, but are also consistent with gender differences in psychological resilience in dealing with stressors leading to girls being less affected regarding their educational outcomes than boys.[34]

#### 5.1.2.2 Analysis by cohort and socio-economic background

In addition, we also investigate heterogeneous effects by cohort, separately for 5th and 9th of primary school and 3rd grade for secondary school and by socio-economic background (income and education). The results can be found in the online appendix, sections C.1 and C.2, respectively.

### 5.2 Student attendance

Attendance is an important input factor in educational production. Lower school attendance as a consequence of exposure to homicides may at least partially explain lower test performance. Aucejo and Romano (2016) found that a reduction in absences at school leads to an increase in both math and reading test scores. We are therefore interested in first understanding whether exposure to homicides around schools affects attendance of students at school, and we use unique individual attendance records of students to whom we have access. Attendance records in São Paulo are available at the bimester. As the ending dates of the bimesters are school specific and these dates are not available centrally, we group the first two and last two bimesters into two semesters.[35] We then calculate the attendance rate of each student for the entire year and in the first and second semesters to use the higher frequency nature of the data. We use the same routine to calculate the explanatory variables. *Homicides (year)* corresponds to the number of homicides within a 25 m radius from school in the school year. *Homicides (1st semester)* and *Homicides (2nd semester)* are the numbers of homicides within a 25 m radius from school in the first and second semesters. In Table 5, we present the regression results of the effect of violence on attendance. In the first column, we present the results for annual attendance records, and in columns (2) and (3), the results for the first and second semesters, respectively.

We find that one additional homicide in the year reduces attendance by approximately 1%.

[34] Evidence from psychology presents mixed results on systematic gender differences in stress resilience, but point to an important distinction between the perception of stressors and coping mechanisms for dealing with stress, leading to an ambiguous effect of stress on objective outcome measures. Day and Livingstone (2003) show that female high school students rated the perceived stressfulness of five hypothetical scenarios higher than male students but were also more likely to seek support. Matud (2004) shows stark differences in the perception of stress, with females subjectively being more stressed than male participants in the study.

[35] We used the official starting and ending dates of each semester provided by the São Paulo State Secretariat of Education.

These results are largely confirmed when examining attendance separately by semester. Each additional homicide around the school in the first semester also reduces attendance in the respective semester by 1%. The coefficients for the second semester exceed the magnitude of the coefficients of the first semester. In the second semester, one additional homicide in the surroundings of the school reduces attendance by about 2% for our preferred specification.[36]

We are also interested in understanding potential heterogeneous effects in line with the previous section. In Table E11, we report the effects on attendance by cohort. We find that attendance at primary school is affected by homicide exposure, whereas we do not find an effect on attendance in secondary school. Table E12 presents the effect of exposure to violence in the school surroundings on attendance in the year and in each semester for boys and girls. We confirm the general pattern across the semesters, with stronger effects in the second semester for both boys and girls. Overall, we find that the effect of homicide exposure on attendance is much more pronounced for boys than for girls, confirming the more pronounced effects for boys in math and language achievement. Finally, we also examine how the effects vary by the socio-economic background of the parents. The results in Table E13 by family income are consistent with the patterns we find for test scores. High income seems to mediate the negative effect of homicides on attendance, and the estimates on absenteeism are much more pronounced for low-income families. When splitting the sample by parental education, we do not find a clear pattern for the effects by family income.

The effects of exposure to homicides on absenteeism are concerning, as low attendance may also hurt achievement. Being an important input factor in educational production, it may also constitute a relevant channel through which violence affects performance on math and Portuguese tests. Alternatively, the effects on attendance and achievement may reflect a general shift away from human capital investments and may therefore be jointly determined.

To determine how much of the results on test scores can be explained by absenteeism alone, we estimate specifications in columns (2) and (5) in Table 1, including student attendance as a control.

[36]This small difference in the outcomes for these semesters could be explained by the dynamic incentives for students to attend over the year. As students can be retained if they fall below a 75% attendance threshold, students may be more prudent regarding their attendance earlier in the school year. Later in the year, when students have more control over their overall yearly attendance, they may be less prudent. We find some evidence for that when comparing the mean attendance rates. In the first semester, attendance is close to 2% higher compared to the second semester. In addition, the law regulating student attendance in São Paulo states that, if a student has accumulated excessive absences, the school must intervene and inform parents, so that they can take measures to remedy the problem. If parents are unsuccessful and the problem persists, the school must notify *Conselho Tutelar*, which is a local legal institution responsible for ensuring the well-being of children and adolescents. This is to attempt to take measures during the year to avoid student repetition due to absences. If students accumulate excessive absences in the first semester, the schools intervene and try to remedy the situation. As a result of the efforts of parents and the schools, the effect in the first semester may decrease. In the second semester, closer to the end of the year, in the event of any negative shock that may affect student attendance, the school may not have time to intervene before the end of the year. Moreover, since it is the end of the year, students may find it harder to catch up with missed classes and potentially miss even more school days.

The results in Table E14 show a decrease of about 17% in the math coefficient and an 11% decrease in the language coefficient. Although the inclusion of an endogenous variable on the right-hand side poses its own concerns, this exercise may explain the role of attendance as an underlying channel explaining the negative effect on achievement. Interestingly, the inclusion of student attendance in either math or Portuguese reduces the coefficient on test scores only minimally. We interpret this as evidence that the reduction in attendance is unlikely to be the main driver of the negative effects on student achievement.

### 5.3 Student and parental aspirations and attitudes towards education

In addition to the objective educational outcomes (test scores and attendance), we have a unique set of self-reported measures available regarding student aspirations, attitudes, and their general perception towards education and school that are not generally available in many other datasets. We can mirror these student-reported variables with information collected from their parents. These outcomes collected in the socio-economic background questionnaire of SARESP put us in a unique position to understand better how exposure to violence affects student aspirations, perceptions about their performance at school, and general attitudes towards school. A similar set of questions answered by their parents allows us to validate the results from a parental perspective.[37] We start by analysing the answers from student-reported aspirations, attitudes, and perception. In addition to their aspiration for post-compulsory education, we are particularly interested in students' general attitudes towards education, their perception of their own performance, and their self-documented home effort towards education (i.e. homework). We use the answers to the binary questions (where agreement with a statement takes a value equal to one, and zero otherwise) as dependent variables and estimate the effect of exposure to homicides using the same specification with the full set of controls as in column (5) of Table 1. We report the estimates separately for boys and girls in Table 6.

We start with the aspiration to continue with post-compulsory education. The question is framed as 'I intend to go to university'. Roughly half of students agree with this statement. We find that exposure to homicides in the school surroundings decreases agreement with this statement for boys by about 3.4%, an 8% reduction in the fraction of boys agreeing with this statement compared

[37]We focus on answers provided by 9th grade students for two reasons. First, the 9th grade socio-economic questionnaire contains the most complete set of answers consistently collected across several waves of SARESP. Second, 9th grade students are at the end of compulsory schooling; hence, their answers regarding their aspirations for post-compulsory education are the most relevant in understanding dropout rates and school transition to secondary education.

to the mean. In contrast, we find the opposite effect for girls. Girls are 4% more likely to agree with the statement when exposed to homicides, but the effect is not significant at conventional levels.

Next, we investigate the effect on self-assessed performance in school. We find that boys are significantly less likely to agree with the statement 'I am a good student'. We find that homicide exposure reduces the propensity to agree with this statement by 14% (compared to a mean of 0.39). We find again, in contrast to boys, the opposite effect on girls. We cannot distinguish whether this reduction for boys is the outcome of reduced effort and willingness to invest in their education, or a change in their perception about their likely performance.

In the next columns, we find that boys were 12% less likely to report that they are *interested in school activities*, while we find no effect on girls. We find a similar pattern for the effects on student effort, measured by their attitude towards homework. We find that boys are less likely report that they *do their homework in time* and are more likely report that they do not do homework at all, while we find the opposite effect for girls.[38]

We also look at additional outcomes related to student attitudes towards school. Boys and girls both less frequently agree that their *school is a nice place*, with a slightly larger coefficient for girls. Boys also less frequently report that they like being at schools, compared to girls as a response to violence exposure, but none of these estimates are statistically significant.

These results provide an intriguing insight into how exposure to homicides changes the aspirations and attitudes towards education differently for boys and girls. When exposed to homicides around school, boys change their attitude towards education and generally display less interest in further education, have a lower perception about their performance at school, and demonstrate lower effort directed towards school, whereas there is no negative effect for girls.

These results from the students are confirmed by the answers from the parent questionnaire. We report these outcomes in Table 7. Parents of boys report that their child is, on average, less interested in school when exposed to violence.[39] They report less frequently that their child likes school (not statistically significant), and less frequently report that their child *is doing well in school* (a reduction by 8% compared to the mean, significant at the 5% level), whereas we do not find any such negative effect of violence on girls reported by the parents. The estimates for girls are either very small or even of the opposite sign, but are not significantly different from zero.[40]

[38] We find that girls are more likely to report doing their homework while watching TV. These estimates are not conditional on doing homework. Indeed, about 9% of boys report not engaging in homework, compared to 5.5% for girls.

[39] This is on a scale from 0 (very negative) to 10 (very positive). The estimated effect for boys corresponds to a reduction by 5% of a standard deviation.

[40] We find no effect on good behaviour at school for boys, but a positive effect for girls; significant at the 10% level.

Parents of boys also less frequently report that their child *studies at home*, confirming the reduction in the student self-reported engagement with homework. The estimated effect corresponds to a 24% reduction, significant at the 10% and 5% levels, for standard and Conley standard errors, respectively. There is no effect for girls. Parents also less frequently report that their child *does their homework in time* and more frequently that their child does *homework while watching TV* for boys. The coefficient for girls is effectively zero for *doing homework in time* and positive and of similar magnitude for *doing homework while watching TV*.

These results confirm the results based on the self-assessment of students. Exposure to violence systematically changes the aspirations and attitudes related to education for boys but does not negatively affect girls. The effects are particularly pronounced for variables more directly measuring current investments and input into education.

To test whether the findings presented in Table 7 on the attitudes of students observed by their parents are simply reflecting a change in the behaviour of their children or reflect a change in the attitude of the parents in response to homicides that differs by sex, we also investigate measures of parental involvement in the schooling of the children. We report the results in Table E15. Across a variety of variables measuring parental involvement, including *helping with their studies at home*, *participation in parent evening*, *talking about school*, and *following the child's homework* we do not find any significant effects of exposure to homicides.[41]

Lastly, we investigate how children report on how involved their parents are with their education. The estimates, reported in columns (9) to (12), on *parents helping with homework* and *asking about homework* show a significant difference between boys and girls. While there is a small and insignificant negative effect for boys on parental help with their homework, the effect is quite pronounced and significant for parents showing an interest by asking about their homework.[42] A caveat of these self-reported measures is that they may reflect both an objective change in parental involvement and a change in the perception of students of their parents' involvement.

Overall, these estimates reveal how the aspirations and attitudes of students assessed by themselves and their parents change differentially for boys and girls in response to homicide exposure. The differences between boys and girls along a number of measures of student attitudes and behaviour related to education are striking and consistent with the differences we report in terms of standardised test scores and attendance, in particular. In Section 6, we revisit the results on

[41]There is some tendency for parents to be less involved in boy's education, for example in *helping children with their studies at home*, for which the difference between boys and girls is quite pronounced.

[42]Boys are 12% less likely to report that their parents ask about homework compared to the mean across all students.

aspirations when investigating the potential underlying transmission channels.

### 5.4 Student progression

In addition to test score results and attendance, we are interested in student progression as additional educational outcome and measure of human capital accumulation. We have these measures for a longer period, 2007 to 2013, and for all cohorts. Because we have information on the addresses of students in municipal schools and can investigate the effect of exposure to homicides not only around schools, but also around their residence, we focus the analysis for these outcomes on students from municipal schools. Linking individual school census records over time enables us to follow students as they progress through their educational careers. We use the fact that we can follow students over time to create the outcome variables grade repetition, dropout rates, and the transition from primary to (non-compulsory) secondary education.[43] Despite efforts to reduce grade repetition and dropout rates, for example through the introduction of automatic grade promotion policies, grade repetition and dropout rates in Brazil remain high. In our sample, 6% of students repeat any given grade, and about 10% of students drop out of school.[44] We are particularly interested in the effect of violence exposure on dropout and school transition, as these have the most profound consequences for the human capital accumulation of students. We focus our attention on the outcome of 'permanent' dropout, which we define as students not re-enrolling in the subsequent period, because it captures better fundamental changes in human capital investment decisions.[45]

In table 8 we present the regression results of the effect of violence on these outcomes for all students in primary and secondary school, by place of exposure. *Panel A* and *Panel B* present the results for exposure around schools and around the residences of students, respectively.[46][47] To boost precision of these estimates, we also combine exposure around schools and residences in *Panel*

[43] Because of the nature of these schooling outcomes, the sample varies by outcome. While we observe students repeating a year for every grade, such that we have the full sample available, we do not create a dropout variable for the final year of secondary school reducing the sample for these cohorts; for the final year of primary year, we create the variable school progression for a single cohort.

[44] This includes students that drop out of school temporarily and re-enrol at a later stage.

[45] Estimates using all formats of dropouts including temporary dropout yield very similar results, and are more pronounced in their magnitude.

[46] For this purpose, we geocoded the addresses linked to the full eight-digit Brazilian postcode (*Código de Endereçamento Postal*). For confidentiality reasons, we are limited to the postcode information of student addresses, different from school addresses and the address of the occurrence of homicides, for which we have the full address details including full street addresses and postcodes. In the urban context of Brazil, these postcodes relate to a relatively small geographic area containing a block of houses. Geocoding these areas returns the centroid of these areas. Because of the measurement error that we introduce by the less precise geocoding, the results are likely subject to attenuation bias, and hence are possibly biased towards zero.

[47] The estimates using exposure at the residence in *Panel B* and the estimates combining school and residence exposure in *Panel C* include both, school and neighbourhood fixed effects.

*C*. *Repetition* is a dummy variable that indicates whether the student attends the same grade in the subsequent year. *Dropout* is a dummy variable that captures whether a student drops out of school during or at the end of the school year and does not re-enrol in the subsequent two years. We are also interested in the transition from primary to secondary school. The variable *school transition* indicates whether students progress to secondary school after completing compulsory education. Roughly 75% of students in our sample continue to secondary school.

We find no effect of exposure to violence on repetition, either around schools, or around the residence, or for the combined exposure at schools or residence.[48] The coefficients are very close to zero and not statistically significant.

Next we investigate student dropout. In *Panel A*, we report a positive effect of exposure to homicides around schools on dropout. Students exposed to a homicide around schools are 2 percentage points, or 15% compared to the mean, more likely to drop out of school, but the coefficient is not statistically significant. For exposure around the residence, we find a positive effect on dropout of 4.3 percentage points, significant at the 1% level of significance. This marks a very strong increase in dropout of about 33% compared to the mean. The effect for the joint exposure in *Panel C*, yields a similar effect size. In line with the exercise for test scores, we also vary the exposure radius for these estimates. In Table E16 we report the estimates for school dropout for a 25 m, 100 m, and 500 m exposure radius by exposure point. Corresponding with the estimates on test scores, we find a reduction in the effect size for the 100 m exposure measure. For the 500 m measure, we no longer find any effect. In Table E17 we also estimate the effects separately for boys and girls. The estimates confirm the differential effect of homicide exposure documented for test scores and attendance. The effect for boys is much more pronounced across the different exposure points, confirming the pattern for test scores and attendance.

The estimates on school transition are small and not statistically significant - in part due to the much smaller number of observations when estimating the effect on school transitions, for either exposure around schools, residence or the combined measure.

The very accentuated and robust results on dropout point to the long-lasting consequences of crime exposure on the human capital accumulation of children in Brazil. Given the stark consequences of dropout, these findings are in line with the negative effect on self-reported aspirations to continue to post-compulsory education, indicating a substantial shift away from further human

[48]We can estimate the effect of exposure to homicides around students' residences only for the school census outcomes because we do not have address information available for the SARESP sample, providing us with test score data and school attendance.

capital investments as a consequence of homicide exposure.

We provide additional results using exposure on the residence-school corridors in the online Appendix B. Within-corridor estimates show that the propensity of students to dropout of school increases substantially after exposure to homicides in the school path. The results are robust to different specifications, such as corridor width and distinct corridors - walking, driving, and public transport.

## 6 Transmission Channels

Exposure to violence may affect educational outcomes through a number of potential channels, where the relative importance of each of these channels likely differs depending on the context. In the case of violence related to conflict, as in Brück et al. (2019) or conflict-like scenarios as in Monteiro and Rocha (2017), the disruption of school supply is likely to affect the quality of the learning environment and hence educational outcomes. The context in our paper differs considerably from the conflict background in Brück et al. (2019) and Monteiro and Rocha (2017) by focussing on day-to-day violence. A potential transmission channel that has received little attention in previous work, is related to the theoretical connection between crime and human capital investments (Soares (2010)). This channel may work through reduced expected life-span similar to Jayachandran and Lleras-Muney (2009) and Oster et al. (2013) or more generally through increased uncertainty about the future. We investigate in this section the evidence for an underlying mechanism based on exposure to violence affecting human capital investments of students in Brazil. We then investigate additional transmission channels previously highlighted in the literature, in particular a channel driven by changes to the supply of schooling. Finally, we also want to test whether the effects we document in this paper are driven by a bereavement effect, where students are affected from the direct loss of a relative or friend. Although these effects may be relevant in itself, the very specific transmission mechanism would make the results difficult to generalise to other contexts.

### 6.1 Human capital investments

An extensive literature has documented the role of life expectancy for the human capital investment decisions of individuals from a theoretical and empirical perspective (Becker (1964), Ben-Porath (1967), Oster et al. (2013)). There also exists a small literature on how changes to life expectancy that differ by sex, such as health and violence shocks, affect investments in education by sex. Jayachandran and Lleras-Muney (2009) use the rapid reduction in maternal mortality linked to the introduction of Sulfa drugs in Sri Lanka to document the effect of life expectancy on human capital investments for girls. Gerardino (2014) showed that, when male-biased violence is high, as measured by homicides rates in Colombian municipalities, boys are less likely to enrol in secondary school relative to girls, possibly due to a reduction in the returns to education.

In Brazil, homicide is a leading cause of death for men up to their mid-twenties and the vast majority of homicide victims are male as shown in Table E1. Exposure to male victims dominated homicides may therefore change the perception of safety of males and females differentially, and hence possibly affect the perceived returns to education for boys more than for girls.[49] In Section 5, we document that the effects on test scores, attendance, and dropout are substantially larger for boys, suggesting that boys react more strongly to the homicide exposure in the school surroundings suggesting a human capital mechanism related to the perceived risk of becoming a victim and in consequence a reduction in life expectancy.

A short-run effect exclusively found for test scores or attendance may be indicative for a mechanism based on short-term effects of exposure to violence related to the stress and trauma. In contrast, any effect working through changing the decisions to invest in human capital should extend beyond a short-term effect on test scores. In Table E5 we showed that the effects on test scores were not driven by homicides close to the test date, and are hence unlikely driven by the short-term stress or short-term effect on the well-being of students. Indeed, we found that the effects were even more pronounced when considering only homicides in the first term, at least 6 month prior to the SARESP test date. Such a longer-term effect is consistent with an underlying channel related to human capital investments, where exposed students alter their investments in their human capital.

Different from Monteiro and Rocha (2017), who find no evidence for the persistence of the effects on test scores, we also document fundamental changes in human capital accumulation that go beyond short-run effects on test scores. Particularly, the effects on dropout indicate fundamental changes in the decision for human capital accumulation of school children with the potential long-term consequences in the labour market. Dropout as an outcome probably best reflects the decision

[49] Alternatively, the effect may work more generally through the increase in uncertainty linked to homicide exposure. There is a related literature documenting the effect of exposure to traumatic experiences and violence on risk preferences (Brown et al. (2019), Cameron and Manisha (2015), Jakiela and Ozier (2019), Moya (2018)) demonstrating that these events generally increase risk aversion, but these papers do not find more pronounced effects on boys. To the contrary, risk aversion of women seems to be more affected by exposure to violence in Kenya (Jakiela and Ozier (2019)), which is inconsistent with our findings being driven by the effects on boys rather than girls.

in further human capital investment, in particular as we are focussing on permanent dropout as an outcome. The effects on dropout are very robust to different exposure points (in the school and residence surroundings, and on the school path) and to different specifications.[50] The magnitude of the effect on dropout is economically very significant – with an increase of about 33% compared to the mean dropout rate – and exceeds substantially the magnitude on outcomes such as test scores and attendance, emphasising the relevance of the effect on student dropout. As for the other outcomes, the effect on dropout is much more accentuated for boys than for girls, consistent with the differential impact of homicide exposure on the human capital accumulation for boys and girls.

The most direct evidence for an underlying mechanism based on human capital channel arises from the findings on the differential effects on aspiration and attitudes for boys and girls presented in Subsection 5.3. Taken together, the pronounced effects for boys across a number of outcomes, the persistence of the effects on dropout, achievement, and attendance, and the differential effect on educational aspiration and attitude presented in this paper, are an indication that exposure to homicides affects the incentives to invest in human capital, and boys being disproportionally affected by this effect.

### 6.2 Teacher attendance and school supply

Having documented that exposure to violence reduces attendance of students at school, as teachers are also exposed to the violence around the school, we test whether exposure to violence affects teacher attendance. We create the teacher attendance rate based on the daily attendance records of teachers and estimate the effect of school exposure on teacher attendance. We report the coefficient in Table E21, column (3). We find no evidence that exposure around schools reduces teacher attendance. The coefficient is extremely small and not statistically significant. Alternatively, we test how much teacher attendance affects the coefficients on test scores estimated in Table 1. We include teacher attendance as an additional control in specifications in columns (3) and (6) of Table 1. The difference in the coefficients when including teacher attendance is minimal (results available upon request).

Furthermore, we investigate whether homicide exposure may lead to other forms of disruption in the school routine, for example through higher teacher or principal turnover. We find no effect on either. These results are contrary to those of Monteiro and Rocha (2017) who stated that the effect

[50] Unfortunately, as the sample for school progression is much smaller, the consistently negative effects for corridor exposure, are not statistically significant.

of exposure to drug battles on educational outcomes is partially caused by teacher absenteeism and turnover but is not unexpected, as their definition of violence exposure is closer to the conflict scenario of the Palestine conflict in the study by Brück et al. (2019).

### 6.3 Bereavement effect

Finally, to check whether the effects we find are driven by grief due to the death of a peer student, a sibling or a friend (who may live in the same neighbourhood, but may not attend the same school) we drop from the explanatory variable all the victims who are 18 years old or younger. We present the results in *Panel B* of Table E5; the specification for all entries follows the most satiated specification of columns (2) and (5) of Table 1. Column (1) shows the effect of homicides around the school including all the victims. In column (2), we exclude all 18-year-old or younger victims. Column (3) considers only male victims in the explanatory variable and column (4) only gunshot victims. Results do not differ in any meaningful way indicating that the negative effects on test scores are not caused by grief for a peer, friend or sibling.[51][52]

## 7 Final Remarks

This paper uses georeferenced data on homicides for Brazil and links these data with measures of school performance to estimate the causal effect of exposure to violence on schooling outcomes and human capital accumulation. We find that students exposed to violence perform worse in math and Portuguese language tests. We find that one additional homicide during the school year leads to a 4.6% of a standard deviation reduction in math and a 5.5% of a standard deviation reduction in language test scores. The results are robust to the inclusion of school-specific time trends and to a battery of robustness checks for selection, spatial correlation, and different specifications. We create indicator variables that allow us to learn about the effects over the entire test score distribution and we find that the effects are more pronounced for students below the median test score. We use very rich information on the student background and find that the effects are particularly pronounced among students from relatively poorer families, possibly suggesting that income works as a buffer against the negative effect of crime. Crucially, we find that the effects across the number

[51] As we constrain the sample of homicide victims to the ages 18 and older, we assume that any peer student, sibling or friend would not be older than 17 for this exercise.

[52] Furthermore, using school attendance records, we find that none of the victims are either a student or a teacher at the school. Details on how we link the student and teacher records with the homicide victim records are available from the authors.

of outcomes are much more pronounced for boys than for girls, indicating substantial differences in the underlying mechanism at work.

Violence around school also affects the attendance of the students at school. Our estimates show that one additional homicide in the year increases absences by around 1%. We nevertheless find that absenteeism can only explain a fraction of the negative effects on the performance measured by standardised tests.

We also document a very substantial effect on dropout of students from school for exposure around the students residence. In addition, we examine exposure to violence in the school path from the residence to school. We use Google APIs and design an algorithm to build corridors along the path line from residences to school and examine exposure to homicides in these corridors. Within-corridor estimates show that the propensity of students to dropout of school increases substantially after exposure to homicides in the school path. The results are robust to different specifications, such as corridor width and distinct corridors - walking, driving, and public transport.

We use extremely rich information on student educational aspirations and attitudes to investigate a number of potential underlying mechanisms. We provide suggestive evidence that exposure to homicides may deteriorate incentives to invest in human capital for boys, who are most likely to be victimised in homicides. We show that the results are not driven by changes in the supply of schooling induced by homicides, for example, by changes in the attendance and turnover of teachers and principals. We also show that the effects are not driven by bereavement for the death of a friend or a teacher.

These results are important to quantify some of the costs of day-to-day violence that go beyond the cost of direct victimisation and have so far being neglected in cost estimates. Improved cost estimates are important for the design of optimal policies targeting crime and violence, including on the prevention of crime. The negative effects we find on measures of school performance, in particular dropout, suggest that violence affects human capital accumulation, possibly leading to long-lasting consequences for the affected children. Since poor neighbourhoods are often more violent, violence is potentially one additional contributor for the socio-economic gradient we observe in many low- and middle-income countries plagued with high crime rates. Because the effects are more concentrated among boys, exposure to violence may also be a contributing factor to the reversed gender gap in education observed in Brazil and other Latin American countries.

## References


Aizer, Anna. 2008. Neighborhood Violence and Urban Youth. *NBER Working Paper Series [National Bureau of Economic Research]* 13773.

Aucejo, Esteban and Teresa F. Romano. 2016. Assessing the Effect of School Days and Absences on Test Score Performance. *Economics of Education Review* 55:70–87.

Becker, Gary. 1964. Human Capital. *University of Chicago Press* .

Ben-Porath, Yoram. 1967. The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy* 75 (4):352–365.

Brollo, Fernanda, Katja Kaufmann, and Eliana La Ferrara. 2019. The Political Economy of Program Enforcement: Evidence From Brazil. *Journal of the European Economic Association* :1–42.

Brown, Ryan, Veronica Montalva, Duncan Thomas, and Andrea Velásquez. 2019. Impact of Violent Crime on Risk Aversion: Evidence from the Mexican Drug War. *Review of Economics and Statistics* Forthcomin.

Brown, Ryan and Andrea Velásquez. 2017. The Effect of Violent Crime on the Human Capital Accumulation of Young Adults. *Journal of Development Economics* 127:1–12.

Brück, Tilman, Michele Di Maio, and Sami H. Miaari. 2019. Learning The Hard Way: The Effect of Violent Conflict on Student Academic Achievement. *Journal of the European Economic Association* :1–36.

Cameron, Lisa and Shah Manisha. 2015. Risk-Taking Behavior in the Wake of Natural Disasters. *Journal of Human Resources* 50 (2):484–515.

Cerqueira, Daniel R. C., Alexandre X.Y. Carvalho, Waldir J. A. Lobão, and Rute I. Rodrigues. 2007. Análise dos Custos e Consequências da Violência no Brasil. *Texto Para Discussão, Ipea* 1284.

Conley, Timothy G. 1999. GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics* 92:1–45.

Cooley-Quille, Michele, Samuel M. Turner, and Deborah C. Beidel. 1995. Emotional Impact of Children's Exposure to Community Violence: A Preliminary Study. *Journal of the American Academy of Child and Adolescent Psychiatry* 34 (10):1362–1368.

Davis-Kean, Pamela E. 2005. The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment. *Journal of Family Psychology* 19 (2):294–304.

Day, Arla L. and Holly A. Livingstone. 2003. Gender Differences in Perceptions of Stressors and Utilization of Social Support Among University Students. *Canadian Journal of Behavioural Science* 35 (2):73–83.

De Fraja, Gianni, Tania Oliveira, and Luisa Zanchi. 2010. Must Try Harder: Evaluating the Role of Effort in Educational Attainment. *Review of Economics and Statistics* 92 (3):577–597.

Dearing, Eric, Holly Kreider, Sandra Simpkins, and Heather Weiss. 2006. Family Involvement in School and Low-income Children's Literacy: Longitudinal Associations between and within Families. *Journal of Educational Psychology* 98 (4):653–664.

Farrell, Albert D. and Steven E. Bruce. 2010. Impact of Exposure to Community Violence on Violent Behavior and Emotional Distress among Urban Adolescents. *Journal of Clinical Child Psychology* 26 (1):2–14.

Flouri, Eirini and Ann Buchanan. 2004. Early Father's and Mother's Involvement and Child's Later Educational Outcomes. *British Journal of Educational Psychology* 74 (2):141–153.

Foureaux Koppensteiner, Martin and Marco Manacorda. 2016. Violence and Birth Outcomes: Evidence from Homicides in Brazil. *Journal of Development Economics* 119:16–33.

Fowler, Patrick J., Carolyn J. Tompsett, Jordan M. Braciszewski, Angela J. Jacques-Tiura, and Boris B. Baltes. 2009. Community Violence: a Meta-analysis on the Effect of Exposure and Mental Health Outcomes of Children and Adolescents. *Development and Psychopathology* 21 (1):227–259.

Gelber, Alexander and Adam Isen. 2013. Children's Schooling and Parents' Behavior: Evidence from the Head Start Impact Study. *Journal of Public Economics* 101 (1):25–38.

Gerardino, Maria Paula. 2014. The Effect of Violence on Educational Gender Gap. *Universitat Pompeu Fabra. Working Paper* .

GormanSmith, Deborah and Patrick Tolan. 1998. The Role of Exposure to Community Violence and Developmental Problems among Inner-city Youth. *Development and Psychopathology* 10:101–116.

Grogger, Jeffrey. 1997. Local Violence and Educational Attainment. *The Journal of Human Resources* 32 (4):659–682.

Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. Parental Education and Parental Time with Children. *Journal of Economic Perspectives* 22 (3):23–46.

Houtenville, Andrew J. and Karen S. Conway. 2008. Parental Effort, School Resources, and Student Achievement. *The Journal of Human Resources* 94:71–89.

Jakiela, Pamela and Owen Ozier. 2019. The Impact of Violence on Individual Risk Preferences: Evidence from a Natural Experiment. *Review of Economics and Statistics* 101 (3):547–559.

Jayachandran, Seema and Adriana Lleras-Muney. 2009. Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines. *Quarterly Journal of Economics* 124(1):349–397.

Matud, M. Pilar. 2004. Gender Differences in Stress and Coping Styles. *Personality and Individual Differences* 37 (7):1401–1415.

McMillen, Daniel, Ignacio Sarmiento-Barbieri, and Ruchi Singh. 2019. Do more eyes on the street reduce Crime? Evidence from Chicago's safe passage program. *Journal of Urban Economics* 110:1–25.

Monteiro, Joana and Rudi Rocha. 2017. Drug Battles and School Achievement: Evidence from Rio de Janeiro's Favelas. *Review of Economics and Statistics* 99 (2):213–228.

Mora, Toni and Josep-Oriol Escardíbul. 2018. Home Environment and Parental Involvement in Homework During Adolescence in Catalonia (Spain). *Youth and Society* 50 (2):183–203.

Moya, Andres. 2018. Violence, Psychological Trauma, and Risk Attitudes: Evidence from Victims of Violence in Colombia. *Journal of Development Economics* 131:15–27.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey. 2013. Limited Life Expectancy, Human Capital and Health Investments. *American Economic Review* 103 (5):1977–2002.

Rockoff, Jonah E. 2004. The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. *American Economic Review* 94 (2):247–252.

Sharkey, Patrick, Ammy E. Schwartz, Ingrid G. Ellen, and Johanna Lacoe. 2014. High Stakes in the Classroom, High Stakes on the Street: The Effects of Community Violence on Students' Standardized Test Performance. *Sociological Science* 1:199–220.

Soares, Rodrigo R. 2010. Welfare Costs of Crime and Common Violence: A Critical Review. *Texto Para Discussão* 581.

Table 1: Effect of exposure to violence around the school on academic achievement

| | *Math proficiency* | | | *Language proficiency* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Homicides* | −2.745 | −2.289 | −3.307 | −2.644 | −2.138 | −2.739 |
| | (2.512) | (1.105)** | (1.350)** | (2.858) | (1.085)** | (1.413)* |
| | [1.773] | [0.850]*** | [0.985]*** | [1.869] | [0.872]** | [0.991]*** |
| Observations | 676,082 | 676,082 | 676,082 | 675,733 | 675,733 | 675,733 |
| Controls | No | Yes | Yes | No | Yes | Yes |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| School x time | No | No | Yes | No | No | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables *Math proficiency* and *Language proficiency* are math and Portuguese standardised test scores normalised at a (250,50) scale. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. **Individual controls** are age, sex and race fixed effects, grade, dummies indicating whether at home the student has access to daily newspaper, magazines, dictionary, novels, poetry and short stories books and encyclopaedias; commuting time from residence to school; age, race, education and employment status of the father and the mother; income and number of people in the house; if parents own the house or rent it; if the house has supply of energy, water, gas, sewage and garbage collection; number of tv's, radios, bathrooms, cars, maids, vacuum cleaners, washing machines, dvd's, fridges, freezers, telephone, computers, cable tv, microwave and internet. **Teacher controls** are sex, age and race of the Portuguese and math teachers. **School controls** are number of staff members, number of school rooms in use and dummies indicating whether the school has computer room, science lab, library, sports court, teachers' room, principal's room, internet connection and if the school offers school meals. **Classroom controls** are share of black students, share of girls and share of students above the appropriate age.

Table 2: Effect of exposure to violence around the school on academic achievement - by neighbourhood crime level

| | *Math proficiency* | | | *Language proficiency* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *All* | *Low* | *High* | *All* | *Low* | *High* |
| *Homicides* | −2.289 | −3.555 | −0.183 | −2.138 | −2.940 | 0.134 |
| | (1.105)** | (1.794)** | (1.509) | (1.085)** | (2.061) | (1.253) |
| | [0.850]*** | [1.407]** | [1.229] | [0.872]** | [1.390]** | [1.142] |
| Observations | 676,082 | 426,653 | 249,429 | 675,733 | 426,709 | 249,024 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. *Low* and *High* refer to crime levels in the neighbourhoods surrounding the schools. We consider a 500 m radius from school and identify schools ever exposed to homicides during the period of 2007 to 2013; then we subtract exposure in the 25 m radius during the period of analysis (2010 to 2013), and classify as *Low* level when the count of homicides is less than or equal to the median and *High* level when it is higher than the median. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables *Math proficiency* and *Language proficiency* are math and Portuguese standardised test scores normalised at a (250,50) scale. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table 3: Effect of exposure to violence around the school on academic achievement - levels of proficiency

| | *Math proficiency* | | | | | *Language proficiency* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| | *Very low* | *Low* | *Median* | *High* | *Very high* | *Very low* | *Low* | *Median* | *High* | *Very high* |
| *Homicides* | 0.010 | 0.014 | −0.022 | −0.012 | −0.006 | 0.010 | 0.019 | −0.014 | −0.008 | −0.005 |
| | (0.010) | (0.011) | (0.010)** | (0.009) | (0.005) | (0.008) | (0.010)* | (0.010) | (0.008) | (0.005) |
| | [0.007] | [0.007]** | [0.008]*** | [0.008] | [0.005] | [0.006]* | [0.007]*** | [0.008]* | [0.007] | [0.004] |
| Observations | 676,082 | 676,082 | 676,082 | 676,082 | 676,082 | 675,733 | 675,733 | 675,733 | 675,733 | 675,733 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. *T*he outcome is a dummy variable taking the value of one for students at or below specific levels of proficiency, and zero otherwise. Very low and Low are students at the 10th and 25th percentile of the test score distribution; and *High* and *Very high* students at the 75th and 90th percentile. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table 4: Effect of exposure to violence around the school on academic achievement - heterogeneous effects by gender

| | *Math proficiency* | | *Language proficiency* | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −2.971 | −1.767 | −2.813 | −1.406 |
| | (1.554)* | (1.075) | (1.326)** | (1.441) |
| | [1.165]** | [0.885]** | [1.048]*** | [1.167] |
| Observations | 335,038 | 341,044 | 334,702 | 341,031 |
| School / time | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables *Math proficiency* and *Language proficiency* are math and Portuguese standardised test scores normalised at a (250,50) scale. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table 5: Effect of exposure to violence around the school on attendance

| | *Attendance (year)* | *Attendance (1st semester)* | *Attendance (2nd semester)* |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Homicides* | −0.010 | | |
| *(year)* | (0.005)** | | |
| | [0.004]** | | |
| *Homicides* | | −0.010 | |
| *(1st semester)* | | (0.004)** | |
| | | [0.004]** | |
| *Homicides* | | | −0.021 |
| *(2nd semester)* | | | (0.005)*** |
| | | | [0.007]*** |
| Mean | 0.879 | 0.888 | 0.870 |
| Observations | 709,386 | 709,386 | 709,386 |
| School / time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Dependent variables are the attendance rates in the year and in each semester. Explanatory variables *Homicides (year)* corresponds to the number of homicides within a 25 m radius from school in the school year; *Homicides (1st semester)* and *Homicides (2nd semester)* are the number of homicides within a 25 m radius from school in the first and second semesters. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table 6: Effect of exposure to violence around school on **student reported** outcomes

| | *I intend to go to university* | | *I am a good student* | | *I like school activities* | | *I do my homework on time* | | *I do my homework watching TV* | | *I do not do homework* | | *My school is a nice place* | | *I like being at school* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −0.034 | 0.024 | −0.054 | 0.020 | −0.035 | −0.013 | −0.007 | 0.030 | −0.004 | 0.010 | 0.003 | −0.006 | −0.005 | −0.018 | −0.003 | −0.002 |
| | (0.022) | (0.016) | (0.029)* | (0.018) | (0.017)** | (0.026) | (0.017) | (0.017)* | (0.020) | (0.016) | (0.009) | (0.011) | (0.027) | (0.043) | (0.019) | (0.023) |
| | [0.017]** | [0.014] | [0.021]*** | [0.013] | [0.014]** | [0.018] | [0.014] | [0.013]** | [0.016] | [0.014] | [0.009] | [0.008] | [0.020] | [0.029] | [0.014] | [0.020] |
| Mean | 0.417 | 0.618 | 0.393 | 0.459 | 0.288 | 0.240 | 0.270 | 0.299 | 0.216 | 0.232 | 0.088 | 0.055 | 0.228 | 0.160 | 0.268 | 0.256 |
| Observations | 97,700 | 104,882 | 99,250 | 106,153 | 98,781 | 105,892 | 96,970 | 104,414 | 96,838 | 104,069 | 97,037 | 104,305 | 99,837 | 106,434 | 98,167 | 105,194 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 9th grade of primary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables are students' answers to a socio-economic questionnaire collected by the school. *I intend to go to university* is a dummy equal to one if the student answers she wants to keep studying, graduate from high school and go to university; and zero otherwise. *I am a good student*, *I like school activities*, *My school is a nice place* and *I like being at school* are a dummies equal to one if students completely agree with the statements and zero otherwise. *I do my homework on time*, *I do my homework watching TV* and *I do not do my homework* are dummies equal to one if the student answers she always does that and zero otherwise. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table 7: Effect of exposure to violence around school on **parent reported** outcomes

| | *Child's interest in school* | | *My child likes school* | | *My child is doing well in school* | | *My child behaves at school* | | *My child studies at home* | | *My child does homework on time* | | *My child does homework watching TV* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −0.124 | 0.011 | 0.008 | 0.023 | −0.032 | 0.012 | −0.020 | 0.004 | −0.036 | −0.005 | −0.024 | 0.002 | 0.022 | 0.020 |
| | (0.084) | (0.079) | (0.020) | (0.021) | (0.017)* | (0.016) | (0.015) | (0.019) | (0.015)** | (0.019) | (0.026) | (0.019) | (0.020) | (0.018) |
| | [0.065]* | [0.071] | [0.017] | [0.017] | [0.014]** | [0.016] | [0.013] | [0.016] | [0.011]*** | [0.015] | [0.019] | [0.015] | [0.015] | [0.016] |
| Mean | 6.880 | 7.482 | 0.403 | 0.369 | 0.390 | 0.473 | 0.461 | 0.567 | 0.148 | 0.214 | 0.287 | 0.313 | 0.355 | 0.410 |
| Observations | 97,219 | 104,126 | 93,003 | 100,922 | 94,510 | 102,346 | 92,355 | 100,574 | 98,733 | 105,632 | 88,648 | 98,032 | 89,638 | 99,878 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 9th grade of primary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables are parents' answers to a socio-economic questionnaire collected by the school. *Child's interest in school* is a rating of child's interest in school by the parents, ranging from 0 -very negative- to 10 -very positive. *My child likes school*, *My child is doing well in school*, *My child behaves at school*, *My child studies at home*, *My child does homework on time* and *My child does homework watching TV* are dummies equal to one if parents completely agree with the statements and zero otherwise. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table 8: Effect of exposure to violence on student progression

*Panel A: Around the school*

| | *Repetition* | *Dropout* | *School transition* |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Homicides* | 0.007 | 0.020 | 0.005 |
| | (0.012) | (0.022) | (0.046) |
| | [0.010] | [0.018] | [0.042] |
| Mean | 0.047 | 0.137 | 0.739 |
| Observations | 2,088,870 | 1,790,101 | 287,304 |
| School / time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

*Panel B: Around the residence*

| | *Repetition* | *Dropout* | *School transition* |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Homicides* | −0.004 | 0.043*** | −0.002 |
| | (0.004) | (0.009) | (0.015) |
| Mean | 0.047 | 0.130 | 0.723 |
| Observations | 1,981,436 | 1,712,188 | 244,302 |
| School/neighb/time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

*Panel C: Around the residence and the school*

| | *Repetition* | *Dropout* | *School transition* |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Homicides* | −0.000 | 0.035*** | 0.012 |
| | (0.005) | (0.011) | (0.018) |
| Mean | 0.047 | 0.130 | 0.723 |
| Observations | 1,981,436 | 1,712,188 | 244,302 |
| School/neighb/time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered (at the school level in Panels A and C and at the neighbourhood level in Panel B) in parentheses.

*Note:* The analysis includes students from the 1st grade of primary school to the 3rd grade of secondary school, over the period of 2007 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school in *Panel A*, from residence in *Panel B* and from school and residence in *Panel C*. Dependent variable *Repetition* is a dummy variable which indicates whether the student has to repeat the same grade as the current year in the coming year. *Dropout* is a dummy variable that captures whether a student drops out of school during or at the end of the school year and does not re-enrol in the subsequent two years. *School progression* indicates if the students in the last grade of primary school progress to secondary school in the subsequent school year, for that reason, regressions for this outcome include only students at the final grade of primary school. Regressions in *Panel A* include time and school fixed effects. Regressions in *Panel B* and *Panel C* include time, neighbourhood and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. **Individual controls** are age, sex and race fixed effects. For a detailed list of school and classroom controls, refer to Table 1 notes.

# Appendix: For Online Publication

## Appendix A Geographic Coordinates and School-residence Corridors

To define the measures of exposure to violence, we geocoded the addresses of the schools, residences, and homicides. For schools, we have the precise address, including street and house number. For the residences, the street and house number are confidential information and cannot be accessed. However, we were granted access to the postcodes and neighbourhoods. In São Paulo, postcodes are quite small units and, in some cases, even more precise than the street names, as streets are typically broken up into several postcodes. For the homicides, we also have the precise location for each case, including street name and (closest) house number.

We used Google Maps API to geocode the addresses. There are five possible geocoding outcomes, which vary depending on the amount of information used in the process: street, neighbourhood, municipality, state, and not found. If the address is geocoded at the street level, it means that the returned result is a precise geocode, for which Google has information down to street address precision. When street-level information is not available, the returned geocoded addresses are approximations, either interpolated between two precise points or the geometric centre of a result, such as a polyline (for example, a street) or polygon (region).

In our analysis, we use only returned addresses geocoded at the street level. Hence, even though we have different levels of information on the addresses of schools, residences, and homicides, the geocoding accuracy level for all these three units is the street level. From the addresses that we geocoded, 96% of the schools and 97% of the residences were geocoded at the street level, and 95% of the homicides in public were geocoded at the street level.

We also used Google to calculate the corridors from residence to school. We used the Google Directions API and calculated path polylines of walking transport mode for each school/residence pair, which we call the *homicide exposure point (HEP)*. For each pair, we went through all the homicide points and calculated the nearest distance between a homicide and that particular polyline. We also calculated walking and straight distances from the residence to school and from the residence to the *HEP*.

To make those calculations feasible and limit the time necessary to run them, we defined some filter rules as follows:

- Define the threshold distance between the homicide points and path polylines to 500 m;
- Ignore walking mode if the straight-line distance is greater than 15 km;
- Define $double - distance = max(straight - linedistance * 2, 500 * 2)$: If $double - distance$ is greater than 100 km, ignore the homicide point outside the circle with a radius of $double - distance/2$ and centre as the middle of the straight line between the school and residence; if $double - distance$ is less

than or equal to 100 km, ignore the homicide point if the straight-line distance between the homicide point and either the school location or residence location is greater than double the distance.

To avoid billions of unnecessary API requests, the straight-line distance calculations, distances along the path of walking distance transport mode polylines, and nearest distance between the homicide points to polylines were all calculated with Google's code without invoking Google APIs. Overall, we used approximately two billion API requests to geocode our data and to generate the corridors for our analysis.

# Appendix B Robustness checks

## B.1 Spatial distribution of homicide exposure

Because schools are often located close to each other in the high-density urban setting of São Paulo, we focus on a 25 m radius around schools as measure of exposure, using the very granular geographic information we have on the addresses of schools and the occurrence of homicides. This minimises potential measurement error from avoiding exposure to the same homicides overlapping across different schools.[53] To test the robustness of the 25 m measure, we also create exposure measures including homicides that are farther away from schools. We expect that the coefficient reduces in size when including homicides that are farther away. This happens for two reasons. First, if we believe that homicides farther away from schools have a weaker effect on students because of the less direct exposure of students, including these homicides will dilute the overall coefficient. Putting it differently, homicides that occur in the very close vicinity of the school likely are much more visible and can be observed by the largest possible fraction of students, whereas homicides farther away are less salient.[54]

Second, once we increase the radius around schools, we find that exposure areas start to overlap more frequently, reducing the signal of the measure. As a robustness check, we therefore estimate regressions in Table 1 using exposure to homicides for larger radii of 100 and 500 m from school. We present the estimates in Table E3. As expected, we find that the coefficients for homicides in a 100 m perimeter are substantially reduced for math and Portuguese language scores. The coefficients are roughly 61% and 58% of the original coefficients, respectively. We lose any effect for exposure at 500 m for math and Portuguese, and the coefficients are close to zero and not statistically significant.[55] Alternatively, we estimate the effect for annuli or 'rings' of different width corresponding to the radii estimated above. As these will not be a weighted average of the original and additional homicides, we expect that the coefficients will drop at a quicker rate when considering homicides in the ring measure. The estimates are presented in Table E4. Indeed, we find that, while the estimates for the annuli of 25-100 m are still negative, the coefficient is reduced at a quicker rate when compared to the 100 m radius measure.

[53] To address potential spatial and serial correlation in our data, we compute Conley (Conley (1999)) standard errors using a weighted average of spatial covariances with a cut point of 25 m. We also computed these standard errors at a 100 m cut-off; the results are unchanged. We find that the standard errors are generally smaller when using the Conley correction, and hence, we are not worried that spatial or serial correlation affects the precision favourably to finding significant estimates when not addressing potential spatial correlation. We report these standard errors in brackets for all presented specifications. In general, spatial standard errors are similar to regular clustered standard errors, confirming that spatial correlation likely plays no major role in our context.

[54] When increasing the radius, homicides that previously were captured in the 25 m radius now define exposure for additional schools but are on average farther away from schools, hence diluting the effect of exposure of the original estimates.

[55] In Figure E6 in the appendix we present further estimates using additional radii of 200, 300 and 400 m, demonstrating that the effects relatively rapidly diminish.

## B.2 Timing of homicide exposure

Although test scores are only available annually, we can still use the high frequency nature of the homicide data to learn about the role of the timing of exposure. First, we use the information on the timing of homicides and the precise test date to learn about whether the results on the performance of students in standardised tests are short lived. To do so, we exclude homicides closer to the test dates from our homicide measure. We present the results in Table E5. To start, we excluded all homicides in the two-week window prior to the test dates. In fact, no homicides occur just prior to the test dates, indicating that the main estimates provided in Table 1 are not caused by short-run effects. This is confirmed by the identical coefficient in columns (1) and (2), and in (6) and (7) for math and Portuguese test scores. In columns (3) and (8), we exclude homicides one month prior to the test. We find very consistent effects compared to the benchmark coefficient; the coefficients for math and Portuguese are even slightly more pronounced. This is also true when excluding all homicides occurring in the second school term. Columns (4) and (9) reveal even more pronounced effects, both for math and Portuguese language performance for homicides occurring more than six months prior to the test date.

This exercise shows that the overall effect is not driven by short-run effects of exposure to homicides just prior to the test date, as would be consistent with effects driven by the short-run stress and a short-run effect on mental well-being of students exposed to homicides.[56] We can also rule out that the effects on test scores are caused by a short-run disruption in the organisation of the tests by homicide exposure around schools or the compositional change of students induced by any short-run effect on the mental well-being of students. The strengthening of the effects for homicides occurring temporally further away from the test date, indicates that any underlying mechanism behind the effects is likely of a longer-term nature. We discuss this in more detail in Section 6.

## B.3 Characteristics of victims

We also use information on the victims and create homicide counts specific to victim characteristics.[57] We use information on the age and sex of the victim and on the cause of death that allows us to categorise homicides by means involving firearms or any other means.[58] We report the effects for these victim characteristics in *Panel B* of Table E5. Compared to the baseline coefficient for math and Portuguese language

[56] Our results contrast in this respect with the findings by Brück et al. (2019) who found that students in schools exposed to conflict-related fatalities during the Second Intifada in the West Bank led to the deterioration in school outcomes in the short-run, which they attribute to the short-term worsening in the students' psychological well-being.

[57] Due to the origin of the data from public health records, namely death records, the information on the characteristics on the crime are relatively limited. For example, we do not have information on the perpetrator in the data or information on the circumstances of the crime, which is sometimes available in crime surveys or police incidence data.

[58] When creating categories of homicide victim characteristics, we are somewhat restrained by relatively small numbers in some categories, which is why we focus on creating relatively coarse main categories. For example, male victims account for roughly 92% of all homicide victims, and homicides by means of gun discharge for roughly 69%. We report descriptive statistics for all available homicide characteristics in Table E1.

reported in columns (1) and (5), we find no pronounced differences by homicide category. In columns (2) and (6), we report the coefficients for victims older than 18. The coefficients are slightly bigger, indicating that the main effects are not driven by homicide victims of similar age to the students in our sample. Next, we report the coefficients only using homicides involving male victims (Columns (3) and (7)). The effects are again slightly bigger, both for math and Portuguese test scores. Finally, we estimate the effect only using victims that were killed by firearm discharge. Again, the coefficients are larger compared to the benchmark, including all homicides, both for math and Portuguese. This finding is consistent with homicides involving the discharge of a firearm being more perceptible by victims or generally being perceived as more serious. Estimating the effects for the largest groups of victims and finding effect sizes in line with the overall effects both for math and language scores reassures us that the effects are not driven by a small number of very specific cases of victims that have an especially large effect. To the contrary, we find evidence that the effects for most generic types of victims might even have slightly more pronounced effects on student achievement in SARESP.

### B.4 Testing for selection in attendance at tests

For a low-stakes test, attendance rates at the SARESP test are high with approximately 87% of students sitting the test. Because of the low-stakes nature of the tests, schools have generally little incentive to manipulate attendance of students at the test, and the scope for selection based on incentives to schools is likely negligible. Despite the high attendance rates, we would like to rule out that attending students are self-selected and that this process is correlated with exposure to homicides. If homicides in the school surroundings affect students' decisions to participate in the test and the propensity to attend differs systematically by student types, this could bias our results.

To test whether students taking SARESP are selected, we start by testing whether violence in the school surroundings affects attendance of students at the math and language tests. For this purpose, we estimate the effect of exposure to homicides in the school surroundings with an indicator on whether a student attended the test, separately for math and for Portuguese. Columns (1) and (8) of Table E6 report the effect on attendance for math and Portuguese, respectively. Both estimated coefficients are small (1.4% and 1.3% compared to the mean) and are not statistically significant. We further test whether the composition of students attending the test differs in any other way. We do this by estimating the effect of homicide exposure on the fraction of boys and girls, white and non-white students, and students from low versus high-income backgrounds. All of the coefficients are small and not significant, and we are therefore confident that self-selection of students into the test does not bias our estimates.[59]

Although within-year transfers across schools are rare, these might lead us to miss selection using the

[59] This is consistent with the fact that the coefficients in Table 1 do not vary across specifications when adding a very large set of socio-economic controls.

above measures for attendance. We therefore test separately whether homicide exposure has an effect on within-year transfers of students. We create an indicator variable taking a value of 1 for students that attended a school at the end of the school year different from the school they were initially registered in at the beginning of the school year. In Table E7, we report the estimates. The coefficients are very small (close to zero) and not statistically different from zero. Taking these results together, we are confident that the estimates are not biased through selected attendance at the SARESP tests.

# Appendix C Additional heterogeneity analysis

## C.1 Analysis by cohort

One additional advantage of the data and identification strategy in this paper is that we can investigate the effect of exposure to homicides for standard outcomes for different age groups. In Table E8, we present the results of the effect of exposure to violence on math and language test scores for each of the three cohorts in our sample: the 5th and 9th grades of primary school and the 3rd grade of secondary school. All specifications include time and school fixed effects and the full set of controls. The coefficients for math are most pronounced for students in the 5th and 9th grades of primary school, for whom an additional homicide in the surroundings of the school during the year implies a reduction of 4.8% and 4% of a standard deviation of math proficiency, respectively. The effect is much smaller and not significant for the 3rd grade of secondary school. We find a very similar pattern for the Portuguese language test scores, with the most pronounced effects for 5th graders and smaller effects for 9th graders and for the final year in secondary school. Splitting the sample by grade nevertheless reduces the precision of the estimates so that, apart from the effects for 9th-grade math test scores, none of the coefficients are separately significant.

## C.2 Analysis by socio-economic status

Next, we use information on parental income and educational background to examine heterogeneous effects by socio-economic status. First, we split the sample by income per capita and classify parents whose family income per capita is less than the median income in each year of the analysis as *low income* and classify others as *high income*. Second, we separately analyse students whose parents have completed, at most, primary school, denoted as *less educated* and students for whom at least one parent has completed at least secondary school, denoted as *more educated*.

In Table E9, we present the results of the effect of violence around schools on test scores separately for each of these categories. All specifications include time and school fixed effects and the full set of controls.[60] Columns (1) and (2) compare math test scores of children in low- and high-income families. We find a much more pronounced and statistically significant negative effect for low-income students, while the effect for high-income students is very close to zero and not statistically significant. We find the same pattern for language proficiency, revealing a similarly stronger effect for students from lower compared to higher income families, as shown in columns (5) and (6).

In columns (3) and (4), we compare the math proficiency of students by the educational background of their parents. Although not significant at conventional significance levels, the results suggest that students whose parents are more educated are more affected in math by exposure to homicides. We observe a similar

[60] This means for the estimates by income we continue to control for educational background of the parents, and for the results by educational background, we continue to control for family income.

pattern for Portuguese test scores, but the differences are less pronounced compared to the socio-economic background. We should emphasise that all estimates in Table E9 include the full set of individual controls (i.e. in columns (1), (2), (5), and (6), we control for the educational background of the parents, and in columns (3), (4), (7), and (8), we control for income).[61]

These results suggest that socio-economic background may have a mediating role. High income seems to provide a buffering mechanism against the harmful effect of exposure to violence. Parents of higher socio-economic status may be better able to shield their children from the negative effects of exposure to violence, possibly through additional safety measures or by providing a sense of security by dropping and picking up their children by car. This is also consistent with the body of literature documenting how parents' socio-economic status may influence children's educational performance through their behaviour and beliefs. In particular, parents of a higher socio-economic status are generally more likely to actively engage in their children's educational process. They are more engaged with teachers, spend more time with their children, and provide more assistance and support for learning at home (Flouri and Buchanan (2004), Davis-Kean (2005), Dearing et al. (2006) Guryan et al. (2008), Houtenville and Conway (2008), De Fraja et al. (2010), Gelber and Isen (2013), Mora and Escardíbul (2018)).

The contrary effects by education are somewhat unexpected. As we simultaneously also control for parental income, these results possibly point to a different mechanism at work, and we can only speculate on the mechanism. More highly educated parents, with everything else equal, possibly may have a better perception of the risks involved when exposed to violence, and in the event of a homicide, they might be more cautious in sending their children to school, hence affecting their children's performance.

[61] We experimented with alternative definitions of high versus low education for parents. Generally, because two individuals are involved, it is much more difficult to define low versus high education households, compared to using income. Alternative classifications (i.e. for high education where both parents have beyond primary education) deliver very similar results.

## Appendix D Exposure on the Residence-School Path

In Section 5.4 our measures for exposure are focused on homicides around schools and the residences of students. Having established an effect of homicide exposure around the school and residence of students on measures of student progression, in particular dropout rates, we would like to investigate further how exposure on the way from the residence to school affects these outcomes. Exposure on the path from the residence to school may be equally salient to exposure at school and the residence, as students would very likely observe the presence of police and emergency services after an occurrence of violent crime, such as a homicide where the victim is assaulted in the public way.[62] To do so, we built corridors as outlined in Section 4 using Google APIs to construct the path based on the shortest distance between the school and residence for each student. Along the walking path line, we construct polygons of 50 m width (25 m to each side of the walking path), which we refer to as *corridors*. We then create a count of the number of homicides occurring within each corridor in a given year. In addition to the walking path, we create alternative corridors based on the shortest driving path and the shortest path using public transport. We focus on educational outcomes presented in the previous section, for which we can build the corridor data. We see these alternative measures simply providing alternative routes to school, that may not necessarily reflect the mode of transport, but rather provide us with robustness checks. Students may for example not use the shortest way to school, but may prefer a slightly longer route, taking them along larger roads etc.

The results are presented in Table E18. *Panel A* presents the outcomes for the walking path. Consistent with the estimates for exposure around schools and the residence, we find very small and insignificant effects on repetition. We confirm the effects for dropout. Starting with the 50 m width walking path, we find that exposure to a homicide leads to a 2 percentage point increase in dropout, a 15% increase compared to the mean, a very similar magnitude compared to exposure at the schools. To boost precision, we widen the corridors. As we ultimately do not know which way students actually take, this will more likely capture exposure to homicides on the path from the residence to school (and vice versa) students actually take. We illustrate this in Figure E5. But doing so may dilute the effect in line with the dilution documented for the school radius. As expected, the effect sizes reduce slightly, but we gain by having more precise estimates. We find that an additional homicide in the 100 m width corridor, leads to an increase in the propensity to drop out of 12%, compared to the baseline.

In *Panels B* and *C*, we investigate the effect for alternative definitions of the residence-school path for driving and public transport. While the Google Maps API uses the respective algorithm to identify the driving and public transport path, we do not regard these necessarily as truly representing different modes

[62] Unfortunately, our violence data, which are based on death certificates, do not contain information on the time of occurrence, which we would have liked to use to concentrate on homicides occurring during likely school commuting periods.

of transport, but see these merely as alternative paths to school.[63] Using these alternative corridors, we find a very consistent positive effect of very similar magnitude compared to the effects for the walking path. For the 50 m width corridor, we find that one additional homicide leads to an increase in the dropout rate of about 2.6 percentage points, a 20% increase compared to the mean. Widening the corridors again reduces the effect size slightly to a 2.1 percentage points. The effects for the public transport path in *Panel C* are very similar, and we find an effect of 2.5 and 2 percentage points for the 50 and 100 m width, respectively. [64] We repeat the exercise focussing on the 50 m width corridors separately for boys and girls in Table E20. We again find much more pronounced effects on dropout for boys than girls, consistent with the above findings on dropout and the findings on other schooling outcomes and students' aspirations and attitudes.

Finally, across the different corridor definitions, we find throughout a negative effect of exposure to homicides in the different corridors (and different widths) on school transition. As we have a substantially smaller sample based on final-year students, for which we can estimate the effect on transitioning to secondary school the year after completing primary education, the estimates are noisier and not statistically significant. These are nevertheless of economically meaningful magnitude, with an additional homicide leading to a reduction of students enrolling in secondary school between half and just over 1 percentage point, a decrease between 1% and 2% compared to the mean progression rate. Increasing the width of the corridor again reduces the magnitude of the estimates in line with expectations.

[63] The driving path may, for example, constitute the safest path to go to school, avoiding shorter but possibly less safe walking paths to school; hence, students may actually walk on this route to school and back.

[64] Alternatively, we estimate the corridor effects including school fixed effects and controlling for corridor length. The results can be found in Table E19. The coefficients are very similar to the corridor fixed effects estimates.

# Appendix E Figures and Tables

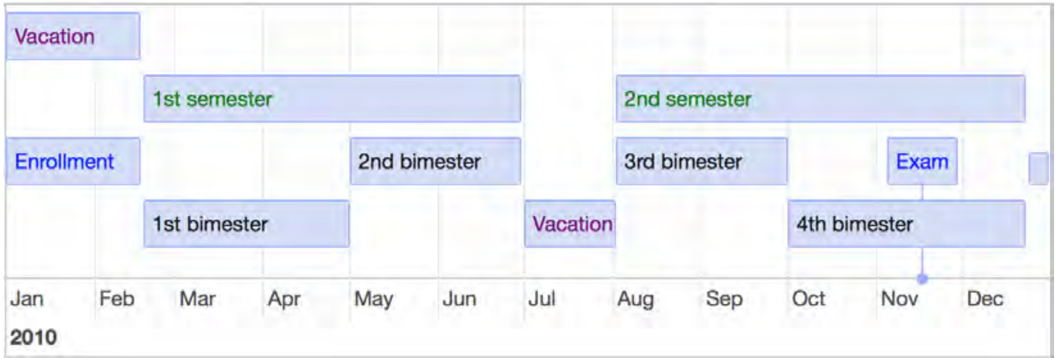



Figure E1: School Calendar in São Paulo

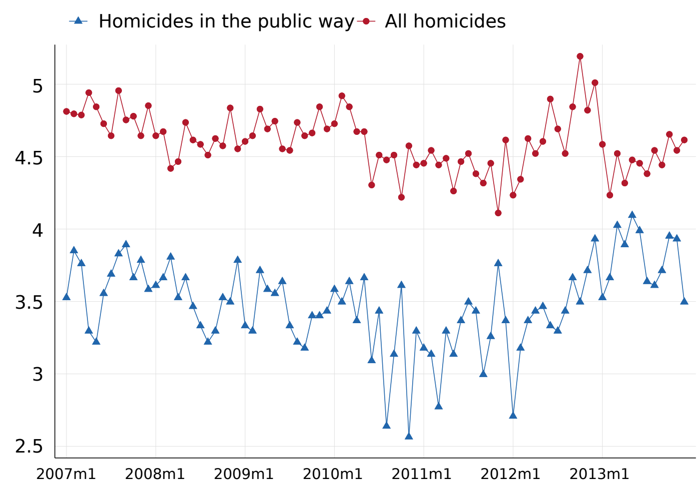
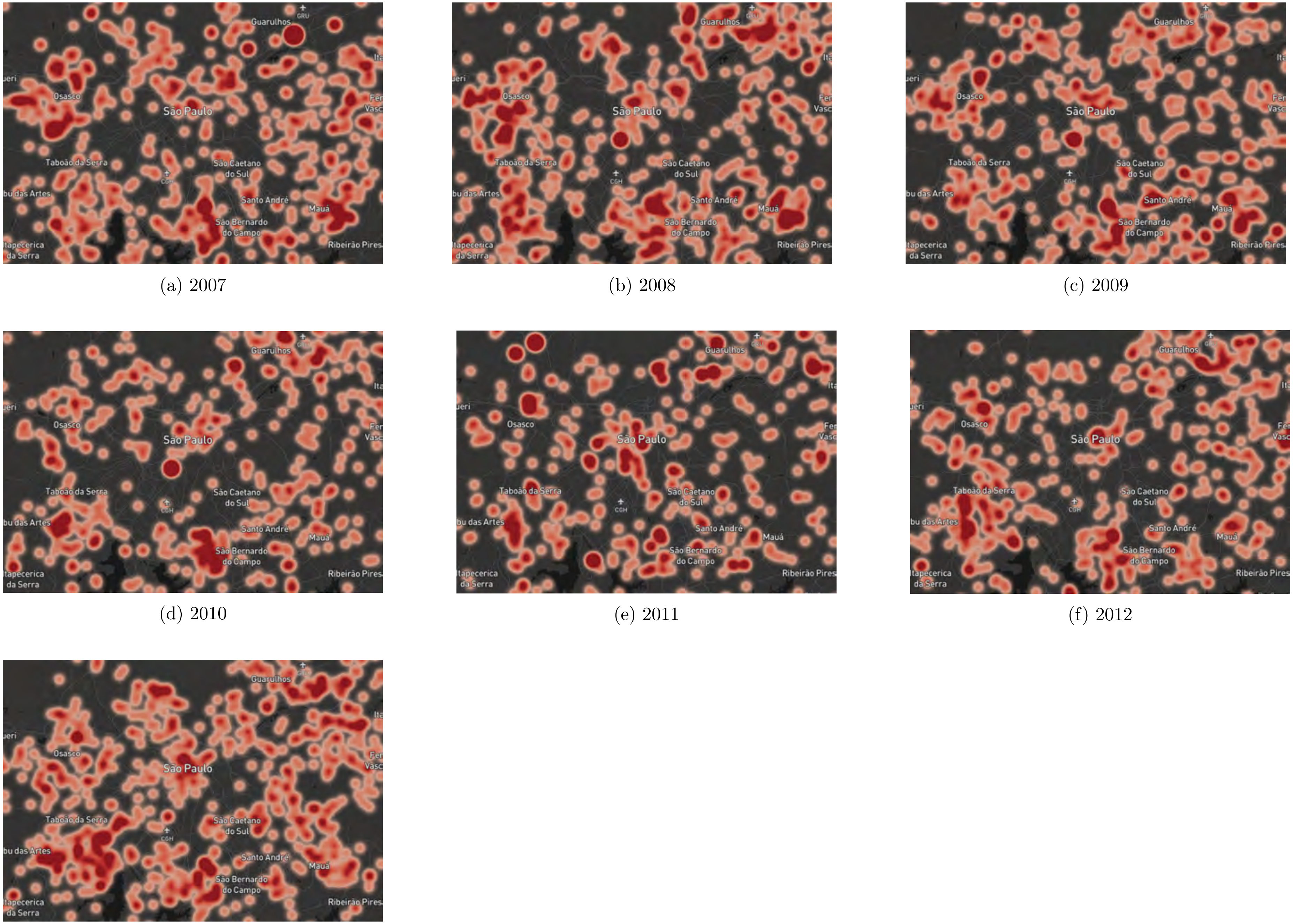



Figure E2: Homicide rates

*Note:* This graph is a comparison between homicides in the public way and the remaining homicide cases, which may occur at a hospital or residence, for example.




(a) 2007

(b) 2008

(c) 2009

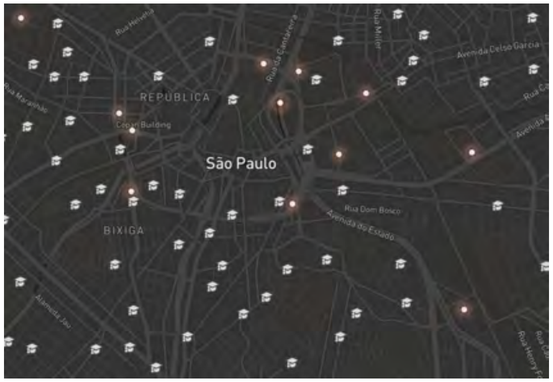
(d) 2010

(e) 2011

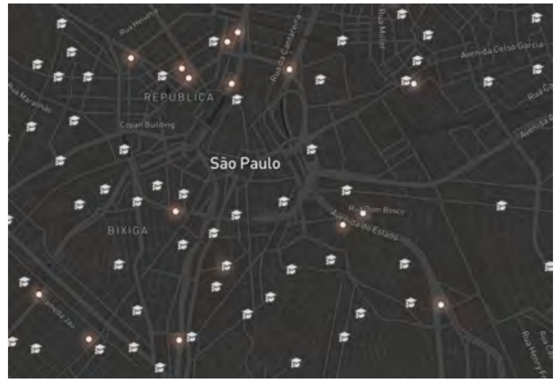
(f) 2012

(g) 2013

Figure E3: Homicides in the public way in São Paulo




(a) 2007




(b) 2008


(c) 2009

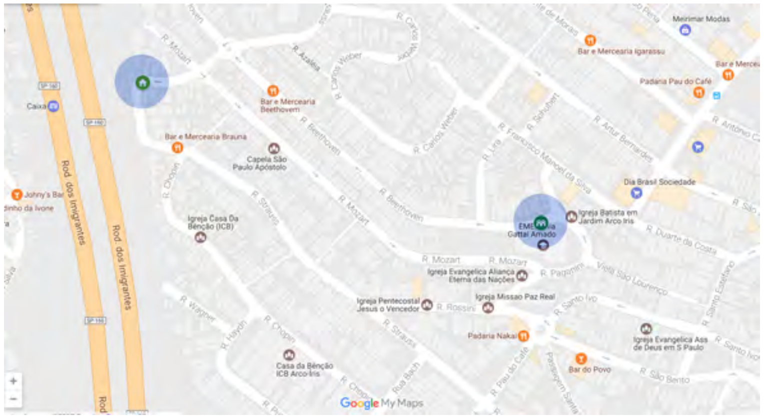

(d) 2010


(e) 2011

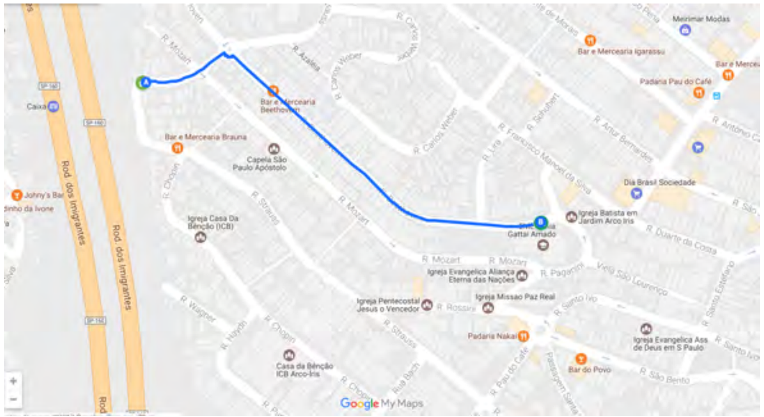

(f) 2012


(g) 2013

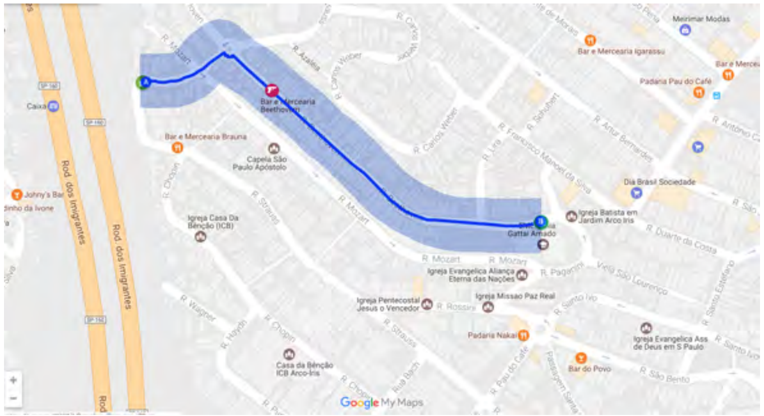
Figure E4: Homicides and schools in a São Paulo neighbourhood



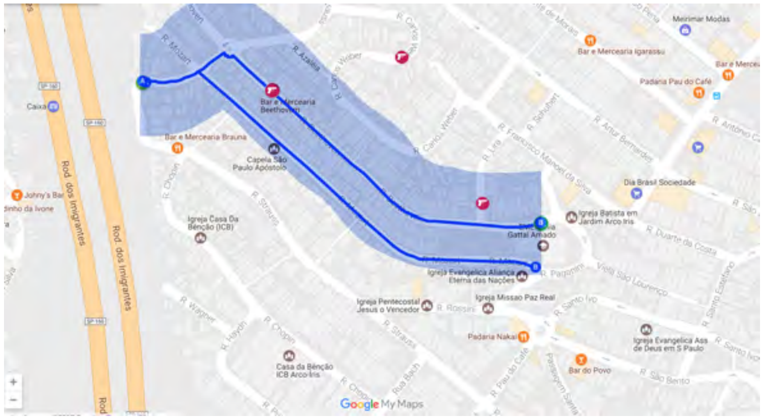
(a) School and residence radius



(b) Shortest walking distance from residence to school



(c) Corridor 1



(d) Corridor 2

Figure E5: Walking path from residence to school - Corridors

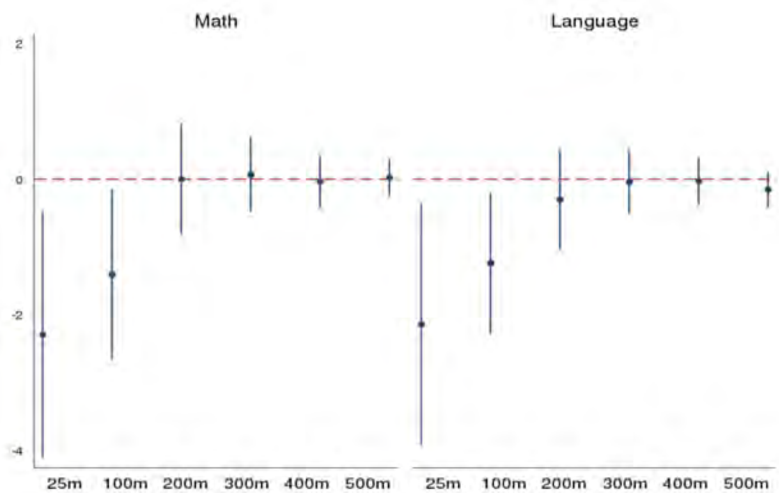



Figure E6: Effect of exposure to violence around the school on academic achievement - 25 m, 100 m, 200 m, 300 m, 400 m and 500 m radii

Table E1: Homicides characteristics

| | ***Homicide victims characteristics*** | |
|---|---|---|
| | Mean | Std.Dev. |
| ***Age*** | | |
| 02-10 | 0.003 | 0.053 |
| 11-15 | 0.021 | 0.143 |
| 16-18 | 0.076 | 0.265 |
| 19-25 | 0.264 | 0.441 |
| 26-30 | 0.191 | 0.394 |
| 31-40 | 0.254 | 0.435 |
| 41-50 | 0.131 | 0.337 |
| 50+ | 0.060 | 0.238 |
| ***Demographics*** | | |
| Male | 0.924 | 0.265 |
| White | 0.420 | 0.494 |
| Black | 0.103 | 0.304 |
| Mixed | 0.453 | 0.498 |
| Single | 0.639 | 0.480 |
| Married | 0.125 | 0.330 |
| Separated | 0.026 | 0.158 |
| ***Education*** | | |
| None | 0.013 | 0.113 |
| 01-03 years | 0.092 | 0.290 |
| 04-07 years | 0.386 | 0.487 |
| 08-11 years | 0.270 | 0.444 |
| 12+ years | 0.033 | 0.179 |
| | ***Homicide characteristics*** | |
| | Number | Percent |
| Assault by gun discharge | 1,709 | 69.190 |
| Assault by sharp object | 273 | 11.053 |
| Assault by blunt object | 256 | 10.364 |
| Assault by bodily force | 137 | 5.547 |
| Assault by other means | 95 | 3.846 |
| Total | 2,470 | 100.000 |

*Note:* The table includes all homicides for which the death occurs in the public way in São Paulo over the period of 2007 to 2013, which were geocoded at the street level.

Table E2: Balancing tests

| | Ever exposed | Never exposed | Diff. | Std. Error |
|---|---|---|---|---|
| ***Students characteristics*** | | | | |
| Age | 13.8946 | 13.3430 | -0.5516 | 0.5070 |
| Female | 0.5148 | 0.5064 | -0.0084 | 0.0544 |
| White | 0.5328 | 0.6133 | 0.0805 | 0.0555 |
| Black | 0.0604 | 0.0497 | -0.0107 | 0.0210 |
| Mixed | 0.3975 | 0.3249 | -0.0726 | 0.0522 |
| Income per capita | 388.4261 | 442.9101 | 54.4841 | 49.6501 |
| Own home | 0.4969 | 0.4309 | -0.0661 | 0.0627 |
| Rent home | 0.5031 | 0.5691 | 0.0661 | 0.0627 |
| Father's education: low | 0.6455 | 0.5584 | -0.0871 | 0.0643 |
| Father's education: mid | 0.2513 | 0.3034 | 0.0522 | 0.0599 |
| Father's education: high | 0.0404 | 0.0766 | 0.0361 | 0.0420 |
| Mother's education: low | 0.5884 | 0.5243 | -0.0640 | 0.0646 |
| Mother's education: mid | 0.3369 | 0.3516 | 0.0146 | 0.0598 |
| Mother's education: high | 0.0484 | 0.0950 | 0.0465 | 0.0486 |
| Father's employment: has a job | 0.4192 | 0.3619 | -0.0574 | 0.0543 |
| Father's employment: has a temp. job | 0.1510 | 0.1260 | -0.0250 | 0.0348 |
| Father's employment: has no job | 0.0336 | 0.0256 | -0.0080 | 0.0117 |
| Mother's employment: has a job | 0.3578 | 0.3104 | -0.0473 | 0.0524 |
| Mother's employment: has a temp. job | 0.1226 | 0.0987 | -0.0240 | 0.0307 |
| Mother's employment: has no job | 0.1225 | 0.0938 | -0.0287 | 0.0273 |
| Travel time from home to school (in min.) | 34.5827 | 34.7566 | 0.1739 | 1.9143 |
| Number of people in the house | 4.4689 | 4.4240 | -0.0449 | 0.1999 |
| Has at home: newspapers | 0.2163 | 0.2328 | 0.0164 | 0.0532 |
| Has at home: magazines | 0.3309 | 0.3485 | 0.0175 | 0.0590 |
| Has at home: dictionary | 0.8762 | 0.8545 | -0.0217 | 0.0437 |
| Has at home: books | 0.8284 | 0.8126 | -0.0158 | 0.0450 |
| Has at home: scientific books | 0.7632 | 0.7490 | -0.0142 | 0.0533 |
| Has at home: water supply | 0.9725 | 0.9685 | -0.0040 | 0.0223 |
| Has at home: sewage supply | 0.8639 | 0.8831 | 0.0192 | 0.0406 |
| Has at home: electricity supply | 0.9638 | 0.9729 | 0.0090 | 0.0185 |
| Has at home: gas supply | 0.2099 | 0.2721 | 0.0622 | 0.0603 |
| Has at home: waste collection | 0.9217 | 0.9307 | 0.0090 | 0.0266 |
| Has at home: television | 0.9646 | 0.9604 | -0.0042 | 0.0244 |
| Has at home: radio | 0.8045 | 0.8122 | 0.0077 | 0.0465 |
| Has at home: bathroom | 0.9092 | 0.9153 | 0.0061 | 0.0311 |
| Has at home: car | 0.4479 | 0.5042 | 0.0562 | 0.0641 |
| Has at home: maid | 0.0749 | 0.1029 | 0.0280 | 0.0425 |
| Has at home: vacuum cleaner | 0.3344 | 0.3802 | 0.0459 | 0.0633 |
| Has at home: washing machine | 0.8548 | 0.8648 | 0.0100 | 0.0395 |
| Has at home: DVD player | 0.8807 | 0.8819 | 0.0012 | 0.0377 |
| Has at home: refrigerator | 0.9276 | 0.9286 | 0.0009 | 0.0295 |
| Has at home: freezer | 0.4956 | 0.4960 | 0.0004 | 0.0626 |
| Has at home: telephone | 0.6769 | 0.6621 | -0.0148 | 0.0592 |
| Has at home: computer | 0.7394 | 0.7492 | 0.0098 | 0.0516 |
| Has at home: cable TV | 0.4797 | 0.5537 | 0.0739 | 0.0622 |
| Has at home: microwave | 0.7670 | 0.7691 | 0.0020 | 0.0497 |
| ***Schools characteristics*** | | | | |
| Computer lab | 0.9250 | 0.9169 | -0.0081 | 0.0549 |
| Science lab | 0.4125 | 0.3839 | -0.0286 | 0.1037 |
| Library | 0.1000 | 0.2061 | 0.1061 | 0.0815 |
| Teachers' room | 0.9500 | 0.9764 | 0.0264 | 0.0300 |
| Principal's room | 1.0000 | 0.9650 | -0.0350 | 0.0373 |
| Sports court | 0.8500 | 0.9401 | 0.0901* | 0.0500 |
| Internet | 0.9875 | 0.9806 | -0.0069 | 0.0214 |
| School meals | 1.0000 | 0.7983 | -0.2017** | 0.0893 |
| Staff members | 89.3000 | 72.8267 | -16.4733** | 7.7087 |
| Number of school rooms in use | 15.5250 | 16.6263 | 1.1013 | 1.8146 |

*Note:* Levels of education are coded as low for parents with up to 8 years of education; mid for parents with secondary school or incomplete high education; and high for parents with complete high education. Employment situation is coded as 'has a job' if parents either have a job, or own a business, or are retired; 'temp. job' if they work independently doing some services, or only do temporary jobs; and 'no job' if they are unemployed.

Table E3: Effect of exposure to violence around the school on academic achievement - 25 m, 100 m and 500 m radii

| | *Math proficiency* | | | *Language proficiency* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *25 meters* | *100 meters* | *500 meters* | *25 meters* | *100 meters* | *500 meters* |
| *Homicides* | −2.289 | −1.403 | 0.029 | −2.138 | −1.233 | −0.149 |
| | (1.105)** | (0.763)* | (0.165) | (1.085)** | (0.631)* | (0.158) |
| | [0.850]*** | [0.602]** | [0.143] | [0.872]** | [0.557]** | [0.140] |
| Observations | 676,082 | 676,082 | 676,082 | 675,733 | 675,733 | 675,733 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m, 100 m and 500 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school in columns (1) and (4), 100 m in columns (2) and (5) and 500 m in columns (3) and (6). Dependent variables *Math proficiency* and *Language proficiency* are math and Portuguese standardised test scores normalised at a (250,50) scale. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E4: Effect of exposure to violence around the school on academic achievement - 'Rings'

| | *Math proficiency* | | | *Language proficiency* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *25 meters* | *25-100 meters* | *100-500 meters* | *25 meters* | *25-100 meters* | *100-500 meters* |
| *Homicides* | −2.289 | −0.978 | 0.099 | −2.138 | −0.805 | −0.100 |
| | (1.105)** | (1.035) | (0.169) | (1.085)** | (0.815) | (0.163) |
| | [0.850]*** | [0.818] | [0.144] | [0.872]** | [0.721] | [0.146] |
| Observations | 676,082 | 676,082 | 676,082 | 675,733 | 675,733 | 675,733 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m, 100 m and 500 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. In columns (1) and (4) explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school; in columns (2) and (5) explanatory variable *Homicides* corresponds to the number of homicides within a 100 m radius from school minus homicides within a 25 m radius from school; in columns (3) and (6) explanatory variable *Homicides* corresponds to the number of homicides within a 500 m radius from school minus homicides within a 100 meter radius from school. Dependent variables *Math proficiency* and *Language proficiency* are math and Portuguese standardised test scores normalised at a (250,50) scale. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E5: Effect of exposure to violence around the school on academic achievement - timing and groups of victims

*Panel A: Homicide timing*

| | *Math proficiency* | | | | | *Language proficiency* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| | *All homicides* | *Excluding last two weeks* | *Excluding last month* | *Excluding 2nd semester* | *Homicides lead* | *All homicides* | *Excluding last two weeks* | *Excluding last month* | *Excluding 2nd semester* | *Homicides lead* |
| *Homicides* | −2.289 | −2.289 | −2.808 | −3.073 | −0.924 | −2.138 | −2.138 | −2.477 | −2.926 | 0.556 |
| | (1.105)** | (1.105)** | (1.106)** | (1.179)*** | (1.565) | (1.085)** | (1.085)** | (1.124)** | (1.264)** | (1.111) |
| | [0.850]*** | [0.850]*** | [0.865]*** | [0.973]*** | [1.365] | [0.872]** | [0.872]** | [0.839]*** | [0.990]*** | [1.045] |
| Observations | 676,082 | 676,082 | 676,082 | 676,082 | 534,837 | 675,733 | 675,733 | 675,733 | 675,733 | 534,573 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel B: Groups of victims*

| | *Math proficiency* | | | | *Language proficiency* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | *All victims* | *18+ yr old victims* | *Male victims* | *Gunshot victims* | *All victims* | *18+ yr old victims* | *Male victims* | *Gunshot victims* |
| *Homicides* | −2.289 | −2.322 | −2.808 | −3.000 | −2.138 | −2.304 | −2.477 | −2.724 |
| | (1.105)** | (1.167)** | (1.106)** | (1.396)** | (1.085)** | (1.136)** | (1.124)** | (1.399)* |
| | [0.850]*** | [0.888]*** | [0.865]*** | [1.053]*** | [0.872]** | [0.915]** | [0.839]*** | [1.009]*** |
| Observations | 676,082 | 676,082 | 676,082 | 676,082 | 675,733 | 675,733 | 675,733 | 675,733 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable Homicides corresponds to the number of homicides within a 25 m radius from school. **In Panel A**, in columns (2) and (7) we exclude from the explanatory variable homicides in the two-week window prior to the test dates; in columns (3) and (8) we exclude from the explanatory variable homicides one month prior to the test; in columns (4) and (9) we exclude from the explanatory variable all homicides occurring in the second school term; in columns (5) and (10) we used homicides lead as explanatory variable. **In Panel B**, in columns (2) and (6) we exclude from the explanatory variable homicides victims younger than 18 years old; in columns (3) and (7) we exclude from the explanatory variable female victims; in columns (4) and (8) we include in the explanatory variable only gunshot victims. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E6: Attendance at Math and Language tests

| | *Attendance at Math test* | | | | | | | *Attendance at Language test* | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
| | *All* | *Boys* | *Girls* | *White* | *Non-white* | *Low income* | *High income* | *All* | *Boys* | *Girls* | *White* | *Non-white* | *Low income* | *High income* |
| *Homicides* | −0.012 | −0.018 | −0.006 | −0.007 | −0.011 | 0.006 | −0.004 | −0.011 | −0.015 | −0.006 | −0.007 | −0.010 | 0.003 | −0.002 |
| | (0.011) | (0.015) | (0.010) | (0.012) | (0.010) | (0.008) | (0.008) | (0.011) | (0.014) | (0.009) | (0.012) | (0.009) | (0.008) | (0.007) |
| | [0.008] | [0.010]* | [0.008] | [0.009] | [0.009] | [0.007] | [0.007] | [0.008] | [0.010] | [0.008] | [0.009] | [0.009] | [0.007] | [0.007] |
| Observations | 777,371 | 388,428 | 388,943 | 271,385 | 207,396 | 191,549 | 220,244 | 777,371 | 388,428 | 388,943 | 271,385 | 207,396 | 191,549 | 220,244 |
| Mean | 0.870 | 0.863 | 0.877 | 0.892 | 0.877 | 0.943 | 0.948 | 0.869 | 0.862 | 0.877 | 0.892 | 0.876 | 0.943 | 0.948 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables *Attendance at Math test* and *Attendance at Language test* indicate whether the student attended the respective exam or not. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E7: Student Mobility

| | Within year transfer | |
|---|---|---|
| | (1) | (2) |
| *Homicides* | −0.001 | 0.001 |
| | (0.002) | (0.003) |
| | [0.003] | [0.003] |
| Mean | 0.016 | 0.016 |
| Observations | 777,371 | 777,371 |
| School / time | Yes | Yes |
| Controls | No | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variable *In year transfer* indicates if the student changes school within the year. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E8: Effect of exposure to violence around the school on academic achievement - heterogeneous effects by cohort

| | Math proficiency | | | Language proficiency | | |
|---|---|---|---|---|---|---|
| | (1) 5th grade (primary school) | (2) 9th grade (primary school) | (3) 3r grade (secondary school) | (4) 5th grade (primary school) | (5) 9th grade (primary school) | (6) 3r grade (secondary school) |
| *Homicides* | −2.407 | −2.032 | −0.890 | −3.435 | −1.733 | 1.072 |
| | (1.649) | (0.934)** | (2.753) | (2.637) | (1.493) | (2.434) |
| | [1.805] | [0.965]** | [2.156] | [2.212] | [1.260] | [2.095] |
| Observations | 266,683 | 298,353 | 111,046 | 266,334 | 298,353 | 111,046 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in 5th and 9th of primary school and 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 meter radius from school. Dependent variables *Math proficiency* and *Language proficiency* are Math and Portuguese standardised test scores normalised at a (250,50) scale. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E9: Effect of exposure to violence around the school on academic achievement - heterogeneous effects by socio-economic status

| | *Math proficiency* | | | | *Language proficiency* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | *Low income* | *High income* | *Less educated* | *More educated* | *Low income* | *High income* | *Less educated* | *More educated* |
| *Homicides* | −2.730 | 0.472 | −0.092 | −1.732 | −3.711 | 0.336 | −0.253 | −0.660 |
| | (1.811) | (1.528) | (1.725) | (1.322) | (1.406)*** | (1.676) | (1.523) | (1.532) |
| | [1.333]** | [1.454] | [1.275] | [1.082] | [1.246]*** | [1.412] | [1.288] | [1.229] |
| Observations | 180,719 | 208,828 | 207,915 | 229,331 | 180,627 | 208,709 | 207,757 | 229,311 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables *Math proficiency* and *Language proficiency* are Math and Portuguese standardised test scores normalised at a (250,50) scale. We coded as *Low income* parents whose income per capita is below the median income in each year and *High income* otherwise. *Less educated* include only cases in which both parents have only primary school and *More educated* cases in which at least one of the parents have more than primary school. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E10: Effect of exposure to violence around the school on assessment of school security by parents

| | *My child is safe at school* | | *My child feels safe at school* | | *My child's school security* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −0.001 | −0.015 | 0.004 | −0.025 | −0.053 | −0.099 |
| | (0.024) | (0.016) | (0.019) | (0.013)** | (0.132) | (0.147) |
| | [0.018] | [0.012] | [0.015] | [0.013]* | [0.093] | [0.115] |
| Mean | 0.289 | 0.222 | 0.326 | 0.258 | 5.171 | 4.981 |
| Observations | 90,091 | 98,212 | 90,842 | 99,032 | 98,206 | 105,150 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 9th grade of primary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables are parents' answers to a socio-economic questionnaire collected by the school. *My child is safe at school* and *My child feels safe at school* are a dummies equal to one if parents completely agree with the statements and zero otherwise. *My child's school security* is a rating of school security by the parents, raging from 0 -very negative- to 10 -very positive. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E11: Effect of exposure to violence around the school on attendance - heterogeneous effects by cohort

| | *Attendance 5th grade (primary school)* | *Attendance 9th grade (primary school)* | *Attendance 3rd grade (secondary school)* |
|---|---|---|---|
| | (1) | (2) | (3) |
| *Homicides* | −0.012 | −0.015 | 0.003 |
| | (0.006)* | (0.006)** | (0.013) |
| | [0.006]** | [0.006]** | [0.009] |
| Mean | 0.915 | 0.854 | 0.866 |
| Observations | 270,865 | 315,760 | 122,761 |
| School / time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Dependent variables are the attendance rates in the year. Explanatory variable *Homicides* correspond to the number of homicides within a 25 m radius from school in the year. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E12: Effect of exposure to violence around the school on attendance - heterogeneous effects by gender

| | *Attendance (year)* | | *Attendance (1st semester)* | | *Attendance (2nd semester)* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −0.015 | −0.007 | | | | |
| *(year)* | (0.006)** | (0.004)* | | | | |
| | [0.005]*** | [0.004]* | | | | |
| *Homicides* | | | −0.017 | −0.008 | | |
| *(1st semester)* | | | (0.007)** | (0.004)** | | |
| | | | [0.006]*** | [0.004]** | | |
| *Homicides* | | | | | −0.031 | −0.012 |
| *(2nd semester)* | | | | | (0.004)*** | (0.005)** |
| | | | | | [0.008]*** | [0.006]* |
| Mean | 0.875 | 0.883 | 0.884 | 0.891 | 0.866 | 0.875 |
| Observations | 353,778 | 355,608 | 353,778 | 355,608 | 353,778 | 355,608 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Dependent variables are the attendance rates in the year and in each semester. Explanatory variables *Homicides (year)*, *Homicides (1st semester)* and *Homicides (2nd semester)* correspond to the number of homicides within a 25 m radius from school in the school year, in the first and in the second semester. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E13: Effect of exposure to violence around the school on attendance - heterogeneous effects by socio-economic status

| | *Attendance (year)* | | | | *Attendance (1st semester)* | | | | *Attendance (2nd semester)* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | *Low income* | *High income* | *Less educated* | *More educated* | *Low income* | *High income* | *Less educated* | *More educated* | *Low income* | *High income* | *Less educated* | *More educated* |
| *Homicides* | −0.005 | 0.001 | −0.002 | −0.005 | | | | | | | | |
| *(year)* | (0.003) | (0.003) | (0.004) | (0.003) | | | | | | | | |
| | [0.003] | [0.003] | [0.003] | [0.003]* | | | | | | | | |
| *Homicides* | | | | | −0.005 | −0.002 | −0.003 | −0.005 | | | | |
| *(1st semester)* | | | | | (0.003)** | (0.003) | (0.003) | (0.004) | | | | |
| | | | | | [0.003]** | [0.002] | [0.003] | [0.003]* | | | | |
| *Homicides* | | | | | | | | | −0.018 | −0.009 | −0.027 | −0.008 |
| *(2nd semester)* | | | | | | | | | (0.004)*** | (0.006) | (0.008)*** | (0.005) |
| | | | | | | | | | [0.005]*** | [0.006] | [0.008]*** | [0.005] |
| Mean | 0.902 | 0.908 | 0.902 | 0.909 | 0.904 | 0.910 | 0.905 | 0.910 | 0.901 | 0.906 | 0.900 | 0.907 |
| Observations | 182,633 | 209,722 | 210,993 | 229,400 | 182,633 | 209,722 | 210,993 | 229,400 | 182,633 | 209,722 | 210,993 | 229,400 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Dependent variables are the attendance rates in each semester and in the year. Explanatory variables *Homicides (year)*, *Homicides (1st semester)* and *Homicides (2nd semester)* correspond to the number of homicides within a 25 m radius from school in the school year, in the first and in the second semester. We coded as *Low income* parents whose income per capita is below the median income in each year and *High income* otherwise. *Less educated* include only cases in which both parents have only primary school and *More educated* cases in which at least one of the parents have more than primary school. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E14: Effect of exposure to violence around the school on academic achievement: the role of students attendance

| | *Math proficiency* | | *Language proficiency* | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Homicides* | −2.128 | −1.776 | −2.204 | −1.958 |
| | (1.194)* | (1.072)* | (1.091)** | (0.999)* |
| | [0.894]** | [0.852]** | [0.897]** | [0.876]** |
| Observations | 641,530 | 641,530 | 641,208 | 641,208 |
| School / time | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes |
| Student attendance | No | Yes | No | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 5th and 9th of primary school and the 3rd grade of secondary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables *Math proficiency* and *Language proficiency* are math and Portuguese standardised test scores normalised at a (250,50) scale. All regressions include time and school fixed effects. Controls include individual characteristics, teachers characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E15: Effect of exposure to violence around the school on parental involvement with education

| | *I help my child studying at home* | | *I participate in my child's parent evening* | | *I talk to my child about school* | | *I follow my child's homework* | | *My parents help me with homework* | | *My parents ask about my homework* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −0.034 | 0.026 | 0.015 | −0.008 | −0.020 | −0.035 | −0.028 | −0.019 | −0.002 | 0.025 | −0.052 | −0.014 |
| | (0.018)* | (0.020) | (0.015) | (0.014) | (0.022) | (0.017)** | (0.016)* | (0.015) | (0.015) | (0.014)* | (0.017)*** | (0.022) |
| | [0.020]* | [0.017] | [0.012] | [0.011] | [0.018] | [0.015]** | [0.014]* | [0.013] | [0.015] | [0.011]** | [0.016]*** | [0.017] |
| Mean | 0.453 | 0.427 | 0.435 | 0.491 | 0.221 | 0.257 | 0.104 | 0.125 | 0.224 | 0.188 | 0.471 | 0.404 |
| Observations | 98,242 | 105,062 | 176,457 | 167,542 | 175,150 | 166,450 | 174,526 | 165,891 | 96,291 | 103,561 | 96,687 | 103,935 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes students in the 9th grade of primary school, over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables are parents and student's answers to a socio-economic questionnaire collected by the school. *I help my child studying at home*, *I participate in my child's parent evening*, *I talk to my child about school*, *I follow my child's homework* are dummies equal to one if parents completely agree with the statements and zero otherwise. *My parents help me with homework* and *My parents ask about my homework* are dummies equal to one if the student answers those situations always happen and zero otherwise. All regressions include time and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. For a detailed list of controls refer to Table 1 notes.

Table E16: Effect of exposure to violence on **dropout** - 25 m, 100 m and 500 m radii

*Panel A: Around the school*

| | (1) | (2) | (3) |
|---|---|---|---|
| | *25 meters* | *100 meters* | *500 meters* |
| *Homicides* | 0.020 | 0.007 | 0.001 |
| | (0.022) | (0.005) | (0.002) |
| | [0.022] | [0.005] | [0.002] |
| Observations | 1,790,101 | 1,790,101 | 1,790,101 |
| School / time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

*Panel B: Around the residence*

| | (1) | (2) | (3) |
|---|---|---|---|
| | *25 meters* | *100 meters* | *500 meters* |
| *Homicides* | 0.043*** | 0.017*** | 0.001 |
| | (0.009) | (0.006) | (0.002) |
| Observations | 1,712,188 | 1,712,188 | 1,712,188 |
| School/neighb/time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

*Panel C: Around the residence and the school*

| | (1) | (2) | (3) |
|---|---|---|---|
| | *25 meters* | *100 meters* | *500 meters* |
| *Homicides* | 0.035*** | 0.007** | 0.000 |
| | (0.011) | (0.003) | (0.001) |
| Observations | 1,712,188 | 1,712,188 | 1,712,188 |
| School/neighb/time | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered (at the school level in Panels A and C and at the neighbourhood level in Panel B) in parentheses.

*Note:* The analysis includes students from the 1st grade of primary school to the 3rd grade of secondary school, over the period of 2007 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school in *Panel A*, from residence in *Panel B* and from school and residence in *Panel C*. Dependent variable *Repetition* is a dummy variable which indicates whether the student has to repeat the same grade as the current year in the coming year. *Dropout* is a dummy variable that captures whether a student drops out of school during or at the end of the school year and does not re-enrol in the subsequent two years. *School progression* indicates if the students in the last grade of primary school progress to secondary school in the subsequent school year, for that reason, regressions for this outcome include only students at the final grade of primary school. Regressions in *Panel A* include time and school fixed effects. Regressions in *Panel B* and *Panel C* include time, neighbourhood and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. **Individual controls** are age, sex and race fixed effects. For a detailed list of school and classroom controls, refer to Table 1 notes.

Table E17: Effect of exposure to violence on student progression - heterogeneous effects by gender

*Panel A: Around the school*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | 0.008 | 0.005 | 0.025 | 0.014 | 0.009 | 0.001 |
| | (0.012) | (0.012) | (0.025) | (0.018) | (0.063) | (0.039) |
| Mean | 0.056 | 0.037 | 0.147 | 0.128 | 0.713 | 0.765 |
| Observations | 1,043,413 | 1,045,307 | 893,519 | 896,577 | 143,759 | 143,545 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel B: Around the residence*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −0.009* | 0.001 | 0.053*** | 0.033*** | −0.005 | 0.002 |
| | (0.005) | (0.004) | (0.011) | (0.008) | (0.014) | (0.024) |
| Mean | 0.056 | 0.037 | 0.140 | 0.120 | 0.694 | 0.752 |
| Observations | 989,203 | 990,912 | 853,649 | 857,184 | 122,697 | 121,343 |
| School/neighb/time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel C: Around the residence and the school*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | −0.002 | 0.002 | 0.043*** | 0.027*** | 0.011 | 0.013 |
| | (0.006) | (0.005) | (0.013) | (0.009) | (0.024) | (0.018) |
| Mean | 0.056 | 0.037 | 0.140 | 0.120 | 0.694 | 0.752 |
| Observations | 989,203 | 990,912 | 853,649 | 857,184 | 122,697 | 121,343 |
| School/neighb/time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered (at the school level in Panels A and C and at the neighbourhood level in Panel B) in parentheses.

*Note:* The analysis includes students from the 1st grade of primary school to the 3rd grade of secondary school, over the period of 2007 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school in *Panel A*, from residence in *Panel B* and from school and residence in *Panel C*. Dependent variable *Repetition* is a dummy variable which indicates whether the student has to repeat the same grade as the current year in the coming year. *Dropout* is a dummy variable that captures whether a student drops out of school during or at the end of the school year and does not re-enrol in the subsequent two years. *School progression* indicates if the students in the last grade of primary school progress to secondary school in the subsequent school year, for that reason, regressions for this outcome include only students at the final grade of primary school. Regressions in *Panel A* include time and school fixed effects. Regressions in *Panel B* and *Panel C* include time, neighbourhood and school fixed effects. Controls include individual characteristics, school characteristics and classroom composition. **Individual controls** are age, sex and race fixed effects. For a detailed list of school and classroom controls, refer to Table 1 notes.

Table E18: Effect of exposure to violence on the residence-school path on student progression

*Panel A: Walking*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *50m width* | *100m width* | *50m width* | *100m width* | *50m width* | *100m width* |
| *Homicides* | 0.003 | −0.000 | 0.020* | 0.015** | −0.006 | −0.003 |
| | (0.004) | (0.003) | (0.010) | (0.007) | (0.017) | (0.012) |
| Mean | 0.047 | 0.047 | 0.130 | 0.130 | 0.724 | 0.724 |
| Observations | 1,876,928 | 1,876,928 | 1,624,079 | 1,624,079 | 231,184 | 231,184 |
| Corridor / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel B: Driving*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *50m width* | *100m width* | *50m width* | *100m width* | *50m width* | *100m width* |
| *Homicides* | 0.005 | 0.000 | 0.026*** | 0.021*** | −0.014 | −0.007 |
| | (0.003) | (0.002) | (0.009) | (0.006) | (0.015) | (0.011) |
| Mean | 0.047 | 0.047 | 0.130 | 0.130 | 0.724 | 0.724 |
| Observations | 1,876,928 | 1,876,928 | 1,624,079 | 1,624,079 | 231,184 | 231,184 |
| Corridor / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel C: Public transport*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *50m width* | *100m width* | *50m width* | *100m width* | *50m width* | *100m width* |
| *Homicides* | 0.004 | −0.001 | 0.025*** | 0.020*** | −0.005 | −0.001 |
| | (0.003) | (0.003) | (0.009) | (0.006) | (0.016) | (0.012) |
| Mean | 0.047 | 0.047 | 0.130 | 0.130 | 0.724 | 0.724 |
| Observations | 1,876,928 | 1,876,928 | 1,624,079 | 1,624,079 | 231,184 | 231,184 |
| Corridor / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the corridor level in parentheses.

*Note:* The analysis includes students from the 1st grade of primary school to the 3rd grade of secondary school over the period of 2007 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within corridors of 50 m and 100 m width. Dependent variable *Repetition* is a dummy variable which indicates whether the student has to repeat the same grade as the current year in the coming year. *Dropout* is a dummy variable that captures whether a student drops out of school during or at the end of the school year and does not re-enrol in the subsequent two years. *School progression* indicates if the students in the last grade of primary school progress to secondary school in the subsequent school year, for that reason, regressions for this outcome include only students at the final grade of primary school. All regressions include time and corridor fixed effects. Controls include individual characteristics, school characteristics and classroom composition. **Individual controls** are age, sex and race fixed effects. For a detailed list of school and classroom controls, refer to Table 1 notes.

Table E19: Effect of exposure to violence on the residence-school path on student progression - school fixed effects controlling for corridor length

*Panel A: Walking*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *50m width* | *100m width* | *50m width* | *100m width* | *50m width* | *100m width* |
| *Homicides* | 0.000 | −0.002 | 0.020*** | 0.015*** | 0.006 | −0.000 |
| | (0.003) | (0.002) | (0.007) | (0.005) | (0.017) | (0.012) |
| Mean | 0.047 | 0.047 | 0.130 | 0.130 | 0.724 | 0.724 |
| Observations | 1,862,503 | 1,862,503 | 1,611,597 | 1,611,597 | 230,129 | 230,129 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel B: Driving*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *50m width* | *100m width* | *50m width* | *100m width* | *50m width* | *100m width* |
| *Homicides* | 0.004 | −0.000 | 0.025*** | 0.021*** | −0.007 | −0.008 |
| | (0.003) | (0.002) | (0.007) | (0.005) | (0.014) | (0.010) |
| Mean | 0.047 | 0.047 | 0.130 | 0.130 | 0.724 | 0.724 |
| Observations | 1,862,502 | 1,862,502 | 1,611,597 | 1,611,597 | 230,129 | 230,129 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel C: Public transport*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *50m width* | *100m width* | *50m width* | *100m width* | *50m width* | *100m width* |
| *Homicides* | 0.001 | −0.002 | 0.023*** | 0.020*** | 0.004 | −0.001 |
| | (0.003) | (0.002) | (0.007) | (0.005) | (0.017) | (0.012) |
| Mean | 0.047 | 0.047 | 0.130 | 0.130 | 0.724 | 0.724 |
| Observations | 1,862,503 | 1,862,503 | 1,611,597 | 1,611,597 | 230,129 | 230,129 |
| School / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses.

*Note:* The analysis includes students from the 1st grade of primary school to the 3rd grade of secondary school over the period of 2007 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within corridors of 50 m and 100 m width. Dependent variable *Repetition* is a dummy variable which indicates whether the student has to repeat the same grade as the current year in the coming year. *Dropout* is a dummy variable that captures whether a student drops out of school during or at the end of the school year and does not re-enrol in the subsequent two years. *School progression* indicates if the students in the last grade of primary school progress to secondary school in the subsequent school year, for that reason, regressions for this outcome include only students at the final grade of primary school. All regressions include time and school fixed effects. Controls include natural log of the calculated path distance, individual characteristics, school characteristics and classroom composition. **Individual controls** are age, sex and race fixed effects. For a detailed list of school and classroom controls, refer to Table 1 notes.

Table E20: Effect of exposure to violence on the residence-school path on student progression - heterogeneous effects by gender

*Panel A: Walking*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | 0.002 | 0.003 | 0.025** | 0.015* | −0.005 | −0.007 |
| | (0.004) | (0.004) | (0.013) | (0.009) | (0.019) | (0.024) |
| Mean | 0.056 | 0.037 | 0.140 | 0.120 | 0.697 | 0.753 |
| Observations | 936,389 | 938,463 | 809,186 | 813,348 | 116,023 | 114,890 |
| Corridor / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel B: Driving*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | 0.005 | 0.005 | 0.033*** | 0.020** | −0.018 | −0.011 |
| | (0.004) | (0.003) | (0.010) | (0.008) | (0.018) | (0.020) |
| Mean | 0.056 | 0.037 | 0.140 | 0.120 | 0.697 | 0.753 |
| Observations | 936,389 | 938,463 | 809,186 | 813,348 | 116,023 | 114,890 |
| Corridor / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

*Panel C: Public transport*

| | *Repetition* | | *Dropout* | | *School transition* | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | *Boys* | *Girls* | *Boys* | *Girls* | *Boys* | *Girls* |
| *Homicides* | 0.004 | 0.004 | 0.031*** | 0.018** | −0.002 | −0.009 |
| | (0.004) | (0.003) | (0.011) | (0.008) | (0.018) | (0.022) |
| Mean | 0.056 | 0.037 | 0.140 | 0.120 | 0.697 | 0.753 |
| Observations | 936,389 | 938,463 | 809,186 | 813,348 | 116,023 | 114,890 |
| Corridor / time | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the corridor level in parentheses.

*Note:* The analysis includes students from the 1st grade of primary school to the 3rd grade of secondary school over the period of 2007 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within corridors of 50 m and 100 m width. Dependent variable *Repetition* is a dummy variable which indicates whether the student has to repeat the same grade as the current year in the coming year. *Dropout* is a dummy variable that captures whether a student drops out of school during or at the end of the school year and does not re-enrol in the subsequent two years. *School progression* indicates if the students in the last grade of primary school progress to secondary school in the subsequent school year, for that reason, regressions for this outcome include only students at the final grade of primary school. All regressions include time and corridor fixed effects. Controls include individual characteristics, school characteristics and classroom composition. **Individual controls** are age, sex and race fixed effects. For a detailed list of school and classroom controls, refer to Table 1 notes.

Table E21: Effect of exposure to violence around school on school supply

| | *Teacher turnover* | *Principal turnover* | *Teacher attendance* |
|---|---|---|---|
| *Homicides* | 0.042 | 0.001 | −0.001 |
| | (0.048) | (0.001) | (0.001) |
| | [0.037] | [0.002] | [0.001] |
| Observations | 92,873 | 2,385 | 124,715 |
| Mean | 0.285 | 0.015 | 0.951 |
| Controls | Yes | Yes | Yes |
| School / time | Yes | Yes | Yes |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the school level in parentheses. Conley standard errors computed at the 25 m cutoff distance in brackets.

*Note:* The analysis includes teachers over the period of 2010 to 2013. Explanatory variable *Homicides* corresponds to the number of homicides within a 25 m radius from school. Dependent variables *Teacher turnover* and *Principal turnover* measure if teachers/principals do not appear in the school system in the following year. *Teacher attendance* is teacher's attendance rate in the school year. All regressions include time and school fixed effects. Controls for regressions on teacher turnover and attendance include individual characteristics and school characteristics. Controls for regressions on principal turnover are school characteristics. **Individual controls** are age, sex and race fixed effects. For a detailed list of school controls, refer to Table 1 notes.