3 Neighborhood Violence and School Achievement: Evidence from Rio de Janeiro's Drug Battles

3.1 Introduction

Violence plagues children across both developed and developing countries, particularly those living in poor urban areas (UNICEF (2006)). There are many reasons to believe that violence has an adverse role in children's educational outcomes, and consequently on their way out of poverty. However, empirical evidence on violence consequences has remained remarkably sparse and mixed. Two main empirical challenges have limited our understanding of the subject. First, neighborhood violence is hardly disentangled from other types of socioeconomic disadvantages that also lead to poor education outcomes (Case & Katz (1991), Aizer (2007), Mayer & Jencks (1989)). Second, it is difficult to characterize and measure local violence. Whenever available, data are usually aggregated at the city level, masking deep variations within cities (Glaeser et al. (1996)), and possibly leading to measurement error of localized violence.

In this paper, we assess whether and how children's educational outcomes are affected by armed conflicts among drug gangs in Rio de Janeiro. Throughout recent decades, several slums scattered across the city have been dominated by heavily-armed drug gangs, which have used the territory to sell drugs and hide from police (Silva et al. (2008), Misse (1999)). Local violence skyrockets when gangs decide to fight each other. Slums are the conflicts epicenter, where risks of life can reach civil war records.[1] Our objective in this paper is to understand whether and how these conflicts affect young children attending the municipal schools located in the proximities of conflict areas. In particular, we examine how student achievement is affected, how students

[1]An assessment made by Extra Newspaper based on police records indicated that 60 percent of the homicides that occurred in the four most violent neighborhoods in Rio de Janeiro metropolitan area (areas 7, 9, 15 and 20) in 2009 were related to drug trade. Source: `http://extra.globo.com/geral/casodepolicia/post.asp?t=morte-obrigatoria-trafico-causa-60-dos-homicidios-nas-zonas-fatais&cod_Post=314033&a=443` This implies that drug trade was responsible for 1,195 deaths only in these areas of the city. Considering the definition of civil war as those internal conflicts that count more than 1,000 battle deaths in a single year (Blattman & Miguel (2010)), this number easily indicates that the drug battles in Rio de Janeiro resemble a civil war.

and teachers respond to these conflicts, and which channels might explain the impact of exposure to violence on educational outcomes. In order to identify where and when armed conflicts take place, we built a novel database on reports about armed conflicts among drug gangs to an anonymous police hotline. We then associate the reports with slums, match this information with educational data by exploring distances between schools and slums, and explore the variation in violence across time and space to identify violence impacts on educational outcomes.

Our results indicate that schools close to areas that experience more variation in armed conflicts over time perform worse in standardized math exams, while no significant effect is found for language exams. Our estimates indicate that a school that experience high levels of violence exposure (percentile 90 of the distribution of violence) scores 0.17 standard deviations less on standardized math tests than a school less exposed to violence (percentile 10). These violent events are also associated with higher grade repetition and dropout rates, particularly for nonwhite students. In terms of mobility across schools, we find no significant effects of violence on students' transfers and new admissions during the school year. Finally, we also discuss the mechanisms underlying these results. We show that violence is associated with higher teacher absenteeism. Violent events during both school and vacation periods affect student achievement, which suggest that violence may have disruptive effects in both the school and the household environments.

This paper contributes to different bodies of literature. Psychologists and psychiatrists have long suggested a positive correlation between children's exposure to local violence and mental health disorders, restricted emotional development, learning problems and truancy (Margolin & Gordis (2000), Fowler et al. (2009), Lynch (2003), Schwartz & Gorman (2003)). In economics, Grogger (1997) shows that violence within schools may reduce the likelihood of high school graduation and the probability of college attendance. Severnini & Firpo (2009) show for a sample of Brazilian schools that students who attend schools which face high levels of school violence usually have lower performance on proficiency tests. However, difficulty in characterizing and measuring violence as well as disentangling its effects from other types of disadvantages has put these evidences into perspective.

The few studies that have tried to overcome identification problems have found little support to a causal relationship between local violence and individual outcomes. Ludwig & Kling (2007) find no support for contagious theories, in which neighborhood violence would induce more crime among individuals. Instead, they find that race segregation may play a more important role in understanding variation across neighborhoods in violent crime than has been thought. Aizer (2007) uses a cross-section survey linking neighborhood violence and children's achievement in math and reading test scores. The author shows that once the family background and other forms of disadvantages are controlled, measures of violence that otherwise negatively affect achievement become mostly insignificant.[2] Thus, while psychologists and psychiatrists have raised concerns over the pervasive role of violence in child development, econometric evidence is scant and still far from a definitive answer for whether the effects of local violence on children's education is a first order policy concern, or if it does only reflect other disadvantages and its side effects.

This paper is also related to two other strands of literature. There is a growing research that aims to understand the effects of extremely violent episodes such as civil conflicts and wars on children's schooling (Akresh & Walque (2008), Shemyakina (ming), León (2009)). However, this literature cannot say much on the mechanisms that explain poor outcomes since it studies major disruptive events, which involve economic and political chaos, and institutional and infra-structure degradation. Finally, this study also contributes to the literature on the social costs of violence (Soares (2006), Lynch & Rasmussen (2001), Hamermesh (1999), Cerqueira et al. (2007)). The results presented in this paper reveal a detrimental effect of violence on individuals' lifetime earnings potential through its impacts on student achievement and accumulation of human capital. Furthermore, our results indicate that violence might create poverty traps, since it makes the way out of poverty more difficult by decreasing student achievement in poor and violent areas.

This paper is organized as follows. Section 2 describes the institutional background, the dynamics of drug gangs conflicts in Rio de Janeiro, and the city's public primary school system. Section 3 presents the data and descriptive statistics of drug conflicts and primary education in Rio de Janeiro between 2004 and 2009. Section 4 discusses a conceptual framework, while section 5 presents our empirical model. Section 6 shows the results, and Section 7 presents robustness checks. Section 8 concludes.

3.2 Institutional Background

[2]There is also a large amount of literature that aims to understand the effects of neighborhoods on individual outcomes (Ludwig et al. (2001), Kling et al. (2005), Kling et al. (2007), Oreopoulos (2003)). In particular, Jacob (2004) explores a natural experiment caused by forced relocation of families due to public housing closure in order to understand the effect of neighborhoods on children's education outcomes. He finds that neighborhoods do not affect education achievement and that family or individual characteristics are much more important to understand differences in educational outcomes than neighborhood influence.

3.2.1 Violence in Rio de Janeiro

Rio de Janeiro is internationally famous for its violence. In 2009, 2,155 people were murdered in the city, which is equivalent to a homicide rate of 32 per 100,000 habitants. This rate is comparable to the ones verified in the most violent cities in United States, such as Detroit (40 murders per 100,000 habitants), Baltimore (37) and Newark (26).[3] This record, already high for international standards, masks striking differences of violence exposure across the city. Poor neighborhoods in the North zone of the city experienced 60.3 deaths for 100,000 inhabitants in 2009, while the South zone rich neighborhoods recorded a homicide rate of approximately 6.6.[4]

Rio de Janeiro violence took off in the early 1980s. This period is marked by the constitution of Comando Vermelho (CV), the first organized drug gang formed in the city, and the entrance of cocaine, which was brought from Bolívia, Peru and Colombia (Dowdney (2003)). Drug dealers relied on the marijuana trade network already established in Rio de Janeiro's slums to sell cocaine. The dominance of slums became crucial to protect the illicit trade. Their geography marked by tiny streets and corners as well as their lawlessness turned slums into an important market for drugs and a strategic place to hide from police (Silva et al. (2008)). The higher profitability of cocaine trade changed drug trade dynamics and led to increasing quarrels among drug members. As a result, some members left Comando Vermelho and created Terceiro Comando (TC) in the late 1980s (Misse (1999)). In the 1990s two other gangs, Amigos dos Amigos (ADA) and Terceiro Comando Puro (TCP), were created by dissidents of the two former gangs. This fractionalization of drug gangs led to more armed conflicts to conquer slums and the increasing militarization of drug gangs (Misse (1997)).

These drug gangs, and more recently the militia[5], are intermittently involved in conflicts where the arsenals employed are similar to the ones found in wars. There is scant research on what triggers these conflicts. Table 3.1 analyzes whether the number of conflicts correlates with slum and neighborhood characteristics. We present coefficients from cross-section regressions of the number of days with conflicts between 2004 and 2009 in

[3]These rates take into account murders and nonnegligent manslaughters, which are defined as the willful (nonnegligent) killing of one human being by another. Source: FBI's Uniform Crime Reporting (UCR) Program.

[4]These rates are from AISP 9 (60.3 deaths per 100,000 inhabitants). The rate of 6.6 deaths per 100,000 inhabitants was registered in AISP 23. Source: http://www.isp.rj.gov.br/ResumoAispDetalhe.asp?cod=200912&mes=Dezembro&ano=2009&tp=Mensal

[5]The militias are groups of policemen and firefighters who pretend to provide security in the neighborhood and charge for this and other 'services" by coercion.

each slum on slum and neighborhood characteristics.[6] Table 3.1 indicates that geographic characteristics such as slum steepness, distance from slum to main roads, slum area and neighborhood population density are good predictors of the number of days with violence, while neighborhood income per capita or income inequality are not associated with the amount of conflicts in slums. This table supports the idea that some slums are more exposed to violence than others because they are strategic places where drug dealers hide from police (a notion supported by the importance of slum steepness in predicting violence) and because of logistical factors (see variable distance to main roads). However, factors that are usually associated with crime, such as income levels and inequality, do not play a role in explaining the prevalence of conflict.

Note that this evidence does not imply that some slums are always embattled, while others are always at peace. Indeed, Figure 3.3, which sketches, for each year, the number of days with conflicts in the ten most violent slums, indicates that levels of violence vary considerably across time.

So, what determines the dynamics of drug conflicts? Newspaper coverage as well as sociologic literature suggest that conflicts are not strategically planned. Baptista et al. (2000) emphasize that Rio de Janeiro's drug gangs are controlled by a group of independent leaders who are inexperienced and young, while Misse (1997) and Souza (2001) argue that gangs do not have a hierarchical structure ruled by a drug baron in the models that we find in Colombia or in the Italian mafia. The reading of the newspaper suggests that these conflicts occur when the unstable power equilibrium among drug gangs is broken by a successful gang overthrow, the imprisonment or release of a gang leader or betrayals. Some fragments of newspaper and blog articles exemplify this argument:

Drug dealers from Morro dos Macacos reobtained the control of three slums in Agua Santa with the support of drug dealers from Rocinha and São Carlos (...) The area was under militia control since last year. The conflict lasted five hours. According to the police department, the invasion was led by Luciano de Oliveira Felipe, known as Cotonete, who is the former slum traffic manager. He was deposed one year ago and was hidden in Morro dos Macacos. (Source: Meia Hora, 6/12/2009)

In addition, drug leaders' release or imprisonment also seems to trigger conflicts as indicated in the following article:

[6]We use in this exercise all the information available at the slum and neighborhood levels

Three people died and eight were wounded after Vila dos Pinheiros invasion by Baixa do Sapateiro drug dealers…The invasion was led by Nei da Conceição Cruz, known as Facão, the main leader of Terceiro Comando Puro (TCP). The conflict began at 10 pm and lasted the whole night. The operation was supported by Matemático.(…) Facão and Matemático left jail last month after winning in Court the right to work outside jail and come back to sleep. Both criminals did not return to jail after the first day under the new sentence. (Source: Meia Hora, 5/31/2009)

In Annex B we transcribe more articles that support our argument. These fragments of newspapers also indicate how violent these events are. People who live in conflict areas and close to them are the most affected. The freedom of movement is drastically affected, the chance of being hit by stray bullets is considerable, and people who are associated with a drug gang can be evicted from their homes or murdered when a new gang assumes control. In addition, these transcriptions show that conflict duration can vary a lot. Conflicts to depose a gang can take hours or days and are usually followed by attempts to reconquer the territory by the former gang. This effort to regain control can occur in the same week or a few months later, depending on how much support the deposed gang can gather from other drug dealers. Therefore, when one conflict begins, it is hard to predict when it is going to be ended.

The impact of these conflicts on city daily routine can be attested with the answers from a victimization survey carried out in 2007. Fear of a stray bullet (60%) and being caught by a gunfight (44%) were mentioned as the violent events of which respondents were most afraid, followed by robberies (37%).[7]

Most of these drug conflicts occur in slums, which does not imply that all slums are controlled by drug gangs and are constantly under conflict. We use slums as the translation for 'favelas", which is defined by the Rio de Janeiro's City Plan as areas characterized by tiny and irregular streets, irregular plot size, poor urban services and irregular settlements.[8] There are 979 slums in Rio de Janeiro according to Instituto Pereira Passos, which concentrate 1.093 million people or 19 percent of the city population (2000 Census data). Figure

[7] This survey was carried out by DATAUFF and interviewed 4,000 people in the Rio de Janeiro metropolitan area. The percentage shown corresponds to answers from people who live in the city of Rio de Janeiro

[8] The definition of favela is given by article 147 of Rio de Janeiro's Plano Diretor (Law number 16/1992): "Para fins de aplicação do Plano Diretor Decenal, favela é a área predominantemente habitacional, caracterizada por ocupação da terra por população de baixa renda, precariedade da infra-estrutura urbana e de serviços públicos, vias estreitas e de alinhamento irregular, lotes de forma e tamanho irregular e construções não licenciadas, em desconformidade com os padrões legais.

3.1 shows the map of the city of Rio de Janeiro with slum borders and indicates that slums are quite widespread around the city.

Although slums are poverty enclaves, not all people in favelas are poor and not all the urban poor live in favelas (Perlman (2010)). Access to urban infrastructure, especially water and electricity distribution, has improved a lot in slums in the last two decades and nowadays are not markedly different from other city areas (Vianna (2008)). But social inequalities are still persistent. In 2007-2008, slum habitants earned 49 percent less than other city habitants and have an average of 3.5 years less education than other city areas (6.4 years of study versus 9.9) (Neri (2010)).

3.2.2 Rio de Janeiro municipal education system

The municipal administration is the main elementary school provider in Rio de Janeiro. The municipal system is one of the largest in Latin America, comprising 1063 elementary schools and 550,000 students.[9] First to fifth graders, which are the focus of our analysis, correspond to 46 percent of the students in the system.[10] There are no school districts in the city and students can choose any school to attend. Some schools have more demand than others, which implies that some students do not end up in their first school choice.[11] The public school network is complemented by the private system, although private coverage is low among poor students. Only 2.5 percent of slum inhabitants attend private schools, while 12.7 percent of other city inhabitants study in the private system (Neri (2010)).

About 36,000 teachers and 13,099 employees work in the municipal school system. All professionals are hired through public exams. Wages are the same across schools but vary with seniority and additional duties. Recently hired teachers are allowed to choose among open placements across different regions, but do not have control over the specific school where they are going to work in the chosen region. There is mobility across schools between years, but it depends on seniority. After three years working in the system, professionals can apply to transfer to another school. Conversations with professionals suggest that some teachers indeed manage to move away from violent areas between years. Within years, however, teachers can only respond to violence shocks with absenteeism and attrition.

[9] Numbers for 2009 gathered at `http://www.rio.rj.gov.br/web/sme/exibeconteudo?article-id=96310`.

[10] We use 1st to 5th graders to refer to students who attend the 1st to 4th grades in the older grade system or the students who are enrolled in the first five years of elementary school according to the new system.

[11] See **?** and **?** for a discussion of the process of registration in public schools in Rio de Janeiro.

Figure 3.1 shows school and slum distribution in the city and indicates that both are widespread. This widespread school distribution and the fact that 98% of children at school age attend school in Rio de Janeiro indicates that school coverage is not a main concern in the city. However, there are several issues related to school quality. An assessment made by the Municipal Secretariat of Education in 2009 showed that 15% of students (28,000) at the 4th, 5th and 6th grades were actually functional illiterates (Prefeitura (2009)). In addition, inequalities across the city are still persistent. Neri (2010) shows that slum inhabitants study less 1 hour and 15 minutes per week compared to other city inhabitants, due to a combination of higher dropout, lower school load and higher absenteeism.

There are several anecdotal evidences of the effects of these conflicts on school routine, mainly among those really close to conflict areas. Problems range from interruption of classes for hours or days, risk of being hit by stray bullets in the way in or out of the school (or sometimes even inside), students' and teachers' emotional disturbs, among others. The headlines from O Globo, the main Rio de Janeiro daily newspaper, exemplify this:

'Teacher is shot by a stray bullet in front of school in Senador Camará", O Globo, March 4th, 2010.

'Gun conflict in Fazendinha left children, who are in their way out of school, in panic", O Globo, June 18th, 2007.

'Boy is shot by a stray bullet inside a school and arrives dead at the hospital", O Globo, July, 16, 2010.

In addition, in a visit to schools located in a highly violent area, we heard several examples of how these conflicts affect school routine. One of the schools did not open for almost an entire month in 2006, when drug gangs were fighting for slum control; teachers and students are intermittently threatened by students connected with drug dealers; and several children, especially the ones who live in the most isolated areas of the slums or the ones with family connections with drug dealers, miss classes or drop out during conflicts. They also mention that children easily identify the bell ringing in the middle of the classes as a signal to leave the classroom and protect themselves from stray bullets in the corridor.

3.3 Data

3.3.1 Violence data

Any understanding of the consequences of Rio de Janeiro's drug conflicts requires finer data about where and when conflicts take place. This is necessary because violence exposure varies dramatically across and within neighborhoods. Official crime data, which is provided by Instituto the Segurança Pública (ISP), cannot track differences in violence exposure since it records information gathered by police stations and then aggregates it for 18 city areas. In addition, ISP does not track information on when and where conflicts happen but only on homicides, which is one of the outcomes of these conflicts. Therefore, we created a novel database for this research based on anonymous reports to Disque-Denúncia, describing that a gun fight occurred in a specific place.

Disque-Denúncia (DD) is a crime hotline that any person can call to report a problem for which she desires the intervention of a public authority. The central was created in 1995 and sits inside the Police Authority of the state of Rio de Janeiro but is managed by an NGO. The calls received by the central are directly forwarded to Civil and Military police, who decide whether and how to respond to each report. All the reports are anonymous and are neither recorded nor tracked. DD works 24 hours a day, 7 days a week and its phone number is broadly disclosed around the city (e.g. on supermarket bags and on buses).

The reports are registered in a database which contains the date, location and description of each event. People call to report any kind of crime and irregularities such as assaults, the location of criminals and bodies, and noise complaints. We got from DD all reports that mention a gun fight among drug gangs between 2003 and 2009 in the city of Rio de Janeiro. We read all reports to guarantee that they described a gunfight and to standardize the addresses provided. The address and the description of the events allow us to associate most of the reports to a specific slum, following the city slum map provided by Instituto Pereira Passos. This procedure generated a list containing all the slums of the city and the dates when a conflict took place. We then aggregated the data per slum and by year by counting the number of days that at least one report of armed conflict was registered in Disque-Denúncia. Annex 1 describes in detail how we built the database.

With the violence measure per slum, we created a measure of violence per school by using GIS tools and considering the distance between each school and each slum. Hence, we defined that the exposure to violence of each school

s at time t is equal to:

$$V_{st} = \sum_j D_{sj} v_{jt}$$

where v_{jt} is the violence level at slum j at time t and D_{sj} is a measure of distance between school s and slum j. Our preferred measure of distance is:

$$D_{sj} = \frac{1}{d_{sj}}$$

where d_{sj} measures the linear distance from school s to slum j closest border. By using this weight, V_{st} considers that each school is exposed to the whole city's violence but gives a higher weight to the violence that occurred closer to the school. Therefore, a particular school is on the top of our violence distribution if it is located closer to one or more violent slums. It can have relatively little exposure to violence if it is surrounded by peaceful slums. We use the logarithm of V_{st} as our main violence measure in the empirical analysis in order to reduce the influence of sharp outliers. We also use as an alternative measure of distance in robustness checks the weight $D_{sj} = 1$ if $d_{sj} \leq x$ meters, which adds the violence of slum j only if it is within x meters from school s, where $x = \{5, 250, 500\}$.

Reports to Disque-Denúncia as a direct measure of violence may raise potential concerns. In section 3.5.2 we provide evidence that Disque-Denúncia reports are indeed a good proxy for violence by comparing it with homicide rates, principals' reports about school violence and by cross-checking with newspaper information.

This research also relies on violence information provided by two newspaper blogs. Plantão de Polícia[12] from Meia Hora newspaper and Casos de Polícia[13] from Extra are to our knowledge the two best information sources of daily violence in Rio de Janeiro's poor areas. We extracted from these blogs the news we transcribed in the Institutional Background section and in the annex. We also used the information provided by these blogs in order to check whether Disque-Denúncia provides a good picture of drug gangs conflicts.

3.3.2 Educational data

In order to determine the impact of drug gang violence on education, we use three databases for educational variables. Students' achievement is measured by Prova Brasil, a national standardized exam applied to all fifth

[12] `http://one.meiahora.com/noticias/cat/plantao-de-policia_26.html`
[13] `http://extra.globo.com/geral/casodepolicia/`

graders in 2005, 2007 and 2009.[14] All students from Rio de Janeiro's schools that had more than 30 students in the fifth grade in 2005 or more than 20 in 2007 and 2009 were supposed to take this exam. The exam is composed of two tests that measure math and language (Portuguese) skills. Unfortunately, the Prova Brasil dataset does not permit students' identification, so we are not able to follow students across time, and we need to rely on score averages at the school level. In addition, students answer a survey about their social-demographic profile, while teachers and principals provide information on their experience and school conditions. In 2007, the principals answered specific questions on violence exposure at school, which we use to compare with our violence data. Prova Brasil dataset is provided by Instituto Anisio Teixeira (INEP). INEP also organizes the Educational Census which provides yearly information on school inputs such as the number of teachers, the number of classrooms, class size, etc. Finally, we use administrative data from Rio de Janeiro's Secretaria Municipal de Educação (SME) from 2004 to 2009. SME gathers information from students' profiles (e.g. date of birth, race, parents' education, religion) when they enter the municipal system and then tracks all their movement within the system. This information includes all public schools each student attended, the grade in which they are enrolled and if and when they transferred between schools. These data allow us to calculate school averages for students' demographics, grade repetition rates, dropout, transfers and new admissions.

3.3.3 Other data

This work relies heavily on geocoded information, which was provided by Instituto Pereira Passos (IPP). Key information is the slum borders, which is based on satellite pictures. This information is not only precise but quite detailed since it defines different slum borders even within large slum areas. As a result, the given definition led to 979 slums (rather than about 300 given by other definitions) which allows us to better localize each violent event. IPP also provides shape files with municipal schools' location, Rio de Janeiro's main roads and neighborhood limits. Based on these shape files, we used GIS tools to calculate the area and population density of Rio de Janeiro's neighborhoods and distances from slums to schools and main roads. In order to understand the determinants of conflicts, we gathered from IPP income per capita, gini index, and population, calculated at neighborhood level based on the 2000

[14]Prova Brasil is also applied to ninth grade students. However, we do not explore this exam because we want to avoid reverse causality. More drug conflicts can lead to more demand for soldiers (older boys), which might impact students' schooling decisions.

IBGE Census. We also obtained information on the slum area for 1999 and 2004. The NASA website provided gridpoints information on Rio de Janeiro's elevation which allowed us to calculate slum steepness. Finally, we got from IPP a list with slum alternative names necessary to match Disque-Denúncia reports to slums.

3.3.4 Summary statistics

Table 3.2 provides Disque-Denúncia descriptive statistics. There were 3,571 reports registered as 'gunfights between drug gangs' from January 1st, 2004 to December 31st, 2009. However, the analysis of the database showed that 444 reports do not describe a gunfight, which led us to exclude them from our analysis.[15] In addition, we exclude another 243 reports that we were not able to associate with a specific slum, leading to a final sample of 2,884 reports.[16] The matching of 92% of the reports to slums confirms the idea that slums are the main conflict battlefield but does not indicate that slum is a synonym of conflict. Table 3.2 shows that about one-third of the slums (289 out of 979) experienced at least one conflict between 2004 and 2009 according to Disque-Denúncia. We refer to this group as violent slums. We see that the average number of reports in violent slums is 1.7 per year or 10 between 2004 and 2009. In our analysis we use the number of days with conflicts in each slum rather than the number of reports in order to deal with the fact that one person can call several times to report the same conflict, leading to striking outliers. We therefore use as our main violence variable the number of days in which there was at least one report about a gunfight. The mean value of this variable in violent slums is 1.4 per year and the standard deviation is 3. The dynamics of these events in the ten most violent slums are exemplified in Figure 3.3. This Figure indicates that violence peaks in different years depending on the slum, which suggests that gunfights are not strategically orchestrated at the city level.

Violence information is associated with the 736 schools (out of 1065) that comprise our sample. These schools are the ones who did Prova Brasil in at least two years between 2005 and 2009. Table 3.3 indicates the proximity of these schools to slums. As already indicated in Figure 3.1, there is no poor supply of schools close to slum areas. We see that 47% of schools are within 250 meters from at least one slum, while 73% are within 500 meters.

[15]The reports that were excluded mention the threat of conflicts among drug gangs, the location of drug dealers, or complement previous information. They are excluded because they do not mention that an armed conflict took place on the specific date.

[16]We were not able to localize the other 243 reports because they do not provide a specific address, or they mention a street that is not inside a slum or close to a slum border.

We focus our analysis on children who attend the first five years of elementary school. The main period of analysis is 2004-2009, the years in which Rio de Janeiro's education administrative data is available. Table 3.4 presents education summary statistics. We show school averages for the whole sample and for violent and non-violent schools. We define as violent schools the ones exposed to violence within 250 meters at any moment between 2004 and 2009. There are 199 violent schools in our sample and 537 schools non-affected by violence within a 250-meter radius. We see that there are marked differences between violent and non-violent schools. Violent schools have worse performance. They do worse in Prova Brasil and have higher failure and dropout rates. However, it is not clear whether the worse performance can be attributed to violence, since these schools enroll more disadvantaged students. Violent schools have a higher share of students with illiterate mothers and fathers, a lower share of whites and less students who live with their parents at home. Interestingly, violent schools have a higher proportion of students who study close to their homes, which indicates that proximity to their household should be an important reason for students to choose these worse performing schools. Violent schools also have more students on average but are not much different from non-violent schools in relation to infrastructure.

The bottom of Table 3.4 also shows principals' reports on the Prova Brasil survey about whether specific events occurred in the school in 2007. Violent schools have a considerably higher incidence of class interruption, students' absence, and drug consumption and trade close to school. These differences not only show that our violence measure correctly indicates schools exposed to drug conflict but also suggest some of the channels through which violence may affect achievement.

In summary, schools exposed to violence are associated with lower achievement. Although this suggests a negative association between violence and achievement, these schools are also attended by students from more disadvantaged households. In the next sections, we discuss the strategies used to disentangle the violence effects on student achievement from other confounding factors.

3.4 Conceptual Framework

In this section, we lay out a simple statistical model that guides our empirical estimations. We first define individuals' exposure to local violence. In a second step, we set a basic model for cognitive achievement in order to highlight the likely channels through which local violence may impact students' performance.

3.4.1 Exposure to violence

Assume $j \in (1, ..., J)$ indexes J slums. We denote the violence level at slum j at moment t as v_{jt}. We assume that v_{jt} depends on two terms. The first term $\overline{v}_j$ is a constant that captures the idea that each slum has an intrinsic level of violence, which depends on neighborhood fixed effects, such as geographic characteristics and strategic position. The second component u_{jt} is an error term uncorrelated with neighborhood characteristics and that follows the normal distribution $u_{jt} \sim N(0, \sigma_j)$. Deviations of u_{jt} might be triggered by events of betrayal and revanchism, responses to threats or imprisonment of gang leaders, and other gang reactions. Therefore,

$$v_{jt} = \overline{v}_j + u_{jt}, \;\; u_{jt} \sim N(0, \sigma_j) \tag{3-1}$$

We assume that child i is exposed to violence v_{jt} depending on the distances between epicenter j to both her household and her school s. Hence, we define students' exposure to local violence by:

$$V_{ist} = \sum_j D_{ij} v_{jt} + \sum_j D_{sj} v_{jt} \tag{3-2}$$

where D_{ij} is the weight that captures the distance between student's i household and slum j, and D_{sj} is the weight that considers the distance between school s and slum j. Replacing (3-1) in (3-2) and rearranging the terms, we have the equation:

$$V_{ist} = (\overline{V}_s + \overline{u}_{st}) + (\overline{V}_i + \overline{u}_{it}) = V_{st} + V_{it} \tag{3-3}$$

Where $\overline{V}_k = \sum_j D_{jk} \overline{v}_j$ and $\overline{u}_{kt} = \sum_j D_{jk} u_{jt}$ for $k = \{s, i\}$. This equation states that students' total exposure to violence is a combination of exposure to violence at school and at home.

3.4.2 Student achievement

Let y_{ist} be a measure of cognitive achievement for child i attending school s at time t. A simple specification of the production function considers that individual's family inputs F_{it} are combined with school resources S_{st} leading to a process of knowledge acquisition as described by the function:

$$y_{ist} = Y\left[F_{it}, S_{st}\right] \tag{3-4}$$

We properly define each of these terms below.

Family inputs

We assume that F_{it} follows the specification

$$F_{it} = F\left[W_{it}, M_{it}(V_{it}), A_{it}(V_{it})\right] \tag{3-5}$$

Or its linearized version

$$F_{it} = \phi_0 + \phi_1 W_{it} + \phi_2 M_{it}(V_{it}) + \phi_3 A_{it}(V_{it}) \tag{3-6}$$

Where W_{it} is an index of the individual's family socioeconomic status at t, $M(\cdot)$ is the individual's mental capacity, and $A(\cdot)$ is the individual's effort to attend classes. We allow $M(\cdot)$ and $A(\cdot)$ to be influenced by an innate cognitive capacity endowment C_i, and also by the individual's exposure to local violence at home V_{it}. Both $M(\cdot)$ and $A(\cdot)$ can be expressed respectively as $M_{it} = \gamma_M C_i + \beta_M V_{it}$ and $A_{it} = \gamma_A C_i + \beta_A V_{it}$. Combining expressions for $M(\cdot)$ and $A(\cdot)$ and equation (3-6), we have:

$$F_{it} = \Phi_0 + \Phi_1 W_{it} + \Phi_2 C_i + \Phi_3\left(V_{it}\right) \tag{3-7}$$

Where Φ_0 collects constant terms, $\Phi_1 = \phi_1$, $\Phi_2 = \phi_2\gamma_M + \phi_3\gamma_A$ and $\Phi_3 = \phi_2\beta_M + \phi_3\beta_A$. In sum, according to the family input channel, shocks of local violence might impact children' outcomes as long as $\Phi_3 \neq 0$. From the standard literature on education production functions we may assume that both ϕ_2 and ϕ_3 are positive parameters (i.e., mental capacity and class attendance are positively correlated with student's achievement). On the other hand, research by psychologists and psychiatrists suggests that $\beta_M < 0$ once exposure to violence might cause mental health disorders and other behavior disturbances. We complement this framework also supposing that $\beta_A < 0$, once local violence may alter students' attendance given the risks it imposes to an individual's movement between household and school. Finally, note that we do not impose any constraints on the relationship between local violence and the family socioeconomic status.

School resources

We assume that S_{st} follows the specification

$$S_{st} = S\left[T_{st}(V_{st}), I_{st}\right] \tag{3-8}$$

Or its linearized version

$$S_{st} = \theta_0 + \theta_1 T_{st}(V_{st}) + \theta_2 I_{st} \tag{3-9}$$

Where T_{st} is, for simplicity, an index measuring the contribution of the representative school's teacher for the student's achievement at time t, and the last term I_{st} represents the school's physical resources. Note that we define T_{st} depending on the level of exposure to violence at the school, while we assume that I_{st} is not affected by violence since these conflicts do not lead to school infrastructure losses. This assumption means that violence effects might occur only through human resources. We assume that T_{st} follows the linear equation

$$T_{st} = \varphi_0 + \varphi_1 HT_{st}(V_{st}) + \varphi_2 MT_{st}(V_{st}) \tag{3-10}$$

Thus, we allow local violence to affect T_{st} via two main channels. First, it may impact teacher's yearly hours of work HT_{st} as defined by expression $HT_{st} = \pi_{1s} + \pi_{ht}V_{st}$. The parameter π_{ht} captures the idea that violence around the school may affect teachers absenteeism, which may even lead to an extreme situation of job attrition during the year, or $HT_{st} = 0$. Second, violence may impact teachers' mental health and their capacity to concentrate and teach. This effect is captured by the expression $MT_{st} = \pi_{2s} + \pi_{mt}V_{st}$. Besides these two channels, local violence might impact the school's human resources via other effects. For instance, the school's staff may react to violent events with more effort in order to alleviate its effects on children. We otherwise omit further extensions in order to simplify the analysis. We combine (3-9) and (3-10) and re-write equation (3-9) as:

$$S_{st} = \Theta_0 + \Theta_1 V_{st} \tag{3-11}$$

where Θ_0 collects constant terms and $\Theta_1 = \theta_1(\varphi_1\pi_{ht} + \varphi_2\pi_{mt})$. We may assume θ_1, φ_1 and φ_2 as positive parameters. Thus, Θ_1 would be a negative parameter as long as local violence negatively affects teachers' yearly hours of work and mental health.

3.5 Empirical Strategy

In this section, we present the statistical models used in our estimations. We first adapt the conceptual framework from the previous section to our empirical setting and to the available data, in order to lay out the baseline regressions. Second, we discuss estimation challenges and potential caveats.

3.5.1 Empirical model

In order to define our empirical model for student achievement, we first linearize equation (3-4):

$$y_{ist} = \rho_0 + \rho_1 F_{it} + \rho_2 S_{ist} + \varepsilon_{ist} \tag{3-12}$$

Where ε_{ist} is an additive measurement error term in test scores. The combination of equation (3-12) with equations (3-7) and (3-11), as well as the introduction of time fixed effects to control for differences among test scores across time, result in the following equation:

$$y_{ist} = \lambda_1 W_{it} + \lambda_2 C_i + \lambda_3 V_{it} + \lambda_4 V_{st} + \mu_s + \vartheta_t + \varepsilon_{ist} \tag{3-13}$$

Where ϑ_t indicates time fixed effects and μ_s represents school fixed effects. This last term collects time-fixed characteristics of the individual's school and its surroundings, such as the school's physical resources, the average quality of human resources and the intrinsic level of exposure to violence in the neighborhood. We are interested in estimates for $\lambda_3 = \rho_1 \Phi_3$ and $\lambda_4 = \rho_2 \Phi_1$, which capture the effects of the exposure to violence at the student's school and household, respectively.

Next we adapt model (3-13) to our empirical setting and to the data available. Our main measure of student achievement is Prova Brasil test scores, which assess math and language skills amongst 5th graders. As mentioned, we are not able to follow students across time and must rely on score averages at the school level. Furthermore, we do not observe individual exposure to violence at home (V_{it}), but only exposure to violence at school (V_{st}). Therefore, we are able to estimate:

$$\overline{y}_{st} = \psi_1 V_{st} + \overline{X}'_{st} \psi_2 + \mu_s + \vartheta_t + \varepsilon^*_{st} \tag{3-14}$$

Where $\overline{y}_{st}$ is the average math or language test scores at school s in year t for 5th graders. The variable of interest is the exposure to violence V_{st}, that varies only at the school level and across time. In our estimations we use the logarithm of this variable to reduce the influence of outliers. The vector $\overline{X}_{st}$ includes controls for student socioeconomic status, which are averaged at the school level and also vary across time. The set of variables in $\overline{X}_{st}$ includes the number of 5th graders per school, average age, share of whites, share of boys and students' mothers education. School and time fixed effects are controlled respectively for μ_s and ϑ_t. Our sample includes the 736 municipal schools that participated in Prova Brasil in at least two years and covers 2005, 2007 and 2009. These are the years in which Prova Brasil was applied. We focus our study on young children from elementary school once they are not subject to soldiering in drug gangs. Thus, we avoid the reverse causality that might occur if conflicts increase the demand for soldiers, which in turn would reduce demand for education.

3.5.2 Empirical challenges

The first caveat underlying our empirical model (3-14) occurs as long as the error term ε^*_{st} includes an omitted variable that captures the students' exposure to violence at home, or $\varepsilon^*_{st} = \varepsilon_{st} + \psi^*\overline{V}_{it}$, where $\overline{V}_{it}$ measures the average exposure to violence at home across students from school s in year t. If $\overline{V}_{it}$ is omitted, and both exposure to violence at school and at home are positively correlated, V_{st} overstates the impact of exposure to violence at school once it also captures effects from $\overline{V}_{it}$. This is a reasonable assumption once students tend to live close to their schools. This concern only affects our understanding of how the total violence effect is disentangled. The positive bias we may find when estimating ψ_1 in model (3-14) due to this limitation should be interpreted as part of the total effect of local shocks of violence.

The second potential caveat relates to individuals' mobility and selection. Achievement test scores are taken at the end of the school year, though any impact violence may have on student's transfers or dropout during the year is of concern. Suppose the error term ε^*_{st} can be re-expressed as $\varepsilon^*_{st} = \varepsilon_{st} + \psi^*\overline{C}_{st}$, where $\overline{C}_{st}$ captures a non observable component of students' cognitive capacity, averaged at the school level. Different arguments may support either a positive or a negative bias on coefficient $\widehat{\psi}_1$ due to the relationship between violence during the school year, students' mobility and $\overline{C}_{st}$. Suppose for instance that violence raises the opportunity costs for less capable students to attend classes. This effect may trigger higher dropout rates among this group of students, leading to higher levels of $\overline{C}_{st}$ at the end of the year, and consequently to a downward bias on $\widehat{\psi}_1$. On the other hand, if violence is associated with low achievement, we can also suppose that more capable students may search for schools in less violent areas during the school year. This effect would bias $\widehat{\psi}_1$ upwards. Thus, the bias direction on $\widehat{\psi}_1$ due to students' mobility is a matter of empirical investigation.

In order to overcome this caveat we proceed as follows. We calculate students' mobility, drop out, retention and attrition using administrative records at the student and school levels. This information allow us to identify whether violence correlates with patterns of grade repetition, dropping out of school and transferring out of school. The administrative records also include students' socioeconomic characteristics, allowing us to examine whether shocks of violence have heterogeneous effects on individuals according to their socioeconomic status. Hence, along with model (3-14), we complement our analysis by estimating the following equations:

$$\overline{z}_{st} = \tau_1 V_{st} + \overline{X}'_{st}\tau_2 + \mu_s + \vartheta_t + \omega_{st} \tag{3-15}$$

Where $\overline{z}_{st}$ indicates the share of students from school s who repeat a grade at the end of the year t or who drop out of or switch to another school in year t. Finally, we complement regressions (3-15) with specifications at the student level for 5th graders which include interactions between violence and socioeconomic characteristics. The linear equations are given:

$$z_{ist} = \kappa_1 V_{st} + (V_{st} * X'_{ist})\kappa_2 + X'_{ist}\kappa_3 + \mu_s + \vartheta_t + \omega_{st} \tag{3-16}$$

Where z_{ist} is a dummy variable at the student level indicating grade repetition, dropping out or transferring. The second term of the right side of equation (3-16) represents interactions between V_{st} and 5th graders' socioeconomic characteristics, included in X_{ist}. The terms μ_s and ϑ_t control, respectively, for school and time fixed effects.[17] We apply this strategy to identifying whether student attrition, retention, as well as drop out rates differ by gender, race, age and mother's education level.

Finally, there is a potential concern related to our measure of violence, since we do not track actual violence, but the number of reports about conflicts. We should stress that the use of such reports to measure violence would be of

[17]Ideally, we should control for student and school fixed effects, but we cannot include both because they are highly correlated. In addition, z_{ist} does not vary much over time. This leads us to choose school, rather than student, fixed effects.

concern to our analysis only if the propensity to report in some neighborhoods changes over time, due to factors also correlated with student outcomes. In order to investigate further this concern, we test several validity checks.

One way to check the validity of Disque-Denúncia data is to cross-check it with official homicide data. Figure 3.4 shows how the number of homicides in the city of Rio de Janeiro and levels of violence documented in Disque-Denúncia reports changed between 2004 and 2009. Note that we are interested in understanding the trends in both variables, rather than comparing levels of violence. The trends in both series are remarkably similar. Both indicate that 2004 was the most violent year; that after 2004, violence declined; but that violence had peaked again by 2009. The largest difference between the two variables occurs in 2006, when a reduction in the number of reports was not followed by a decrease in the number of homicides. Figure 3.5 shows the yearly correlation between the number of homicides and the number of days with conflicts, aggregated per AISP (the city division used by the police department). We observe that in all years, there is a strong correlation between the two measures, which vary from 0.48 in 2004 to 0.74 in 2006 and 2007. Therefore, comparing the number of homicides to Disque-Denúncia shows that Disque-Denúncia data provide a reasonable picture of variations in violence across time and space.

In addition, comparing Disque-Denúncia data with homicide data offers clues to whether the propensity to report changes over time. Figure 3.5 indicates that each AISP consistently tends to be situated above or below the prediction lines, suggesting that a regional propensity to over or under-report is constant over time. Table 3.5 formalizes this finding by showing the actual and predicted homicide based on the number of days with reports in each AISP and year, and on whether the region over or under-reported violence each year. This exercise indicates that 11 AISPs always over-report violence, i.e., have a predicted homicide level greater than the actual number, while five AISPs always under-report. Only AISPs 14 and 31 demonstrate changes in their propensity to report over time.[18] These two AISPs are located in Rio de Janeiro's Western Zone, a region which was marked during the period under analysis by increasing in militia dominance. There is evidence that the militia intimidates the local population (see Cano & Ioot (2008) and Soares et al. (2010)) which can change the propensity to report conflicts. Although it is not

[18]AISP 14 includes the following neighborhoods: Anchieta, Guadalupe, Parque Anchieta, Ricardo de Albuquerque, Campo dos Afonsos, Deodoro, Jardim Sulacap, Magalhães Bastos, Realengo, Vila Militar, Bangu, Gericinó, Padre Miguel and Senador Camará. AISP 31 includes Barra da Tijuca, Camorim, Grumari, Itanhangá, Joá, Recreio dos Bandeirantes, Vargem Grande and Vargem Pequena

clear what the militia's effect on student outcomes might be, we deal with this concern in the robustness check by excluding Rio de Janeiro's Western Zone from the sample.

Another way to validate our data is to compare it with principals' answers about student exposure to violence at school. In Prova Brasil survey, principals were asked about whether specific events had happened at their schools in 2007. In Table 3.4 we showed how the means of these variables differed between violent and non-violent schools. Table 3.6 shows the correlation between our main measure of violence and principals' answers, after controlling for principals' characteristics (e.g. how long they have been on the job, their education), students' average characteristics (e.g. share of blacks, share of females) and school inputs. Each column has a different dependent variable that indicates whether the cited event happened in a given school in 2007. Table 3.6 indicates that Disque-Denúncia measure is correlated with violent events associated with drug traffic and consumption outside schools but not with robbery and violence inside schools. This corroborates our argument that we explore drug traffic violence and not other types of violence, such as robberies and assaults.

A final way of checking the validity of our measure of violence is to compare Disque-Denúncia reports with newspapers coverage. We read all the news about violence in Casos de Policia and Plantão de Policia blogs in 2009. All the conflicts among drug gangs that were mentioned in the blogs corresponded to at least one report in our database. But Disque-Denúncia offered a much more complete picture of gang conflicts because it cited events that were not covered by the newspapers. Unfortunately, the information provided in the blogs gives few details of events, which made it difficult to draw systematic comparisons between newspaper and Disque-Denúncia coverage.

3.6 Results

This section presents our empirical results as follows. We first present the baseline estimations for achievement test scores. Second, we show results for students' mobility, dropout and grade repetition. Finally, we discuss mechanisms that could explain the relationship between violence and student achievement.

3.6.1 Achievement test scores

The baseline results for model (3-14) are presented in Table (3.7). Panel A presents results for achievement test scores in math, while Panel B shows results for language. The first column follows a random effect specification, which includes only year fixed effects. We see that violence is negatively correlated with students' achievement in both panels, and highly robust in Panel A. In the second column we include controls for students' socioeconomic composition. As a result, we see that significance vanishes in both Panels. This result is consistent with cross-section empirical evidence supporting the importance of socioeconomic disadvantages in explaining children's outcomes vis-a-vis local violence effects. However, the third column shows that once school fixed effects are introduced in our longitudinal analysis, the negative effects of violence stand out, particularly in Panel A. We provide evidence below that this result is robust to more flexible measures of violence, controlling for outliers and selecting for different samples. The estimated coefficient in column 3, Panel A, indicates that moving from the bottom decile (p10) to the top decile (p90) of the distribution of violence across schools is associated with a score 1.76 points lower in math exams. This implies that a school in a highly violent area scores 0.17 standard deviations less on standardized math tests than schools in relatively peaceful areas.[19]

The differences between math and language results are of particular interest. Coefficients and significance are higher in Panel A, much in line with the idea that violence effects on math scores are relatively more pervasive once math learning is more demanding in terms of concentration and instruction.

3.6.2 Mobility, dropout and grade repetition

As mentioned above, any impact violence may have on students' patterns of mobility and dropout during the year is of concern once achievement test scores are taken at the end of the school year. Yet, lower math scores can be driven by changes in the composition of students during the year. For instance, if more capable students move from schools in violent areas towards less violent areas during the year, we overstate the estimated effect of violence. Furthermore, besides this potential bias, the relationship between violence and mobility patterns is also of policy concern by its own, particularly if violence hit more disruptively the probability of dropping out of those children from more disadvantaged households.

Table 3.8 presents results for model (3-15) for 5th grade students. The first columns shows that exposure to violence raises the rate of grade repetition for 5th graders. This result reinforces the evidence from Table 3.7 once grade repetition can be thought as an alternative measure for student achievement - yet endogeneity must be of concern given that schools may adjust repetition criterion in response to local violence. The estimated coefficient for 5th graders'

[19]We use as reference math score standard deviation across schools in the base year (2005).

repetition rates indicates that students from schools in the top decile of the distribution of violence have a failure rate 2 percentage points higher than students from schools situated in the bottom decile (p10). The magnitude of this effect represents 23% of the sample average.

Along with the results from Table 3.7, the result from column 1 may otherwise show an upward biased coefficient if more capable students leave the schools under violent conflicts during the year. Thus, in the remaining columns we examine dropout and mobility patterns. In the second column, we see that exposure to violence is significantly associated with higher rates of dropout for 5th graders. The estimated coefficient indicates that dropout rates are 3 percentage points higher in a school localized in a violent area (p90 of the distribution of violence) when compared to schools from relatively peaceful areas (p10). The magnitude of this effect represents 10% of the sample average. Columns 3 and 4 show, respectively, that neither the share of students transferred out of school nor the share of new admissions during the year are significantly associated with violence. Thus, Table 3.8 shows that violence affects students selection only by increasing the number of those students who drop out. This effect may have a selective impact on the remaining pool of students at the end of the year depending on the type of students that are more likely to dropout due to violent events. In order to examine further this question, we run model (3-16) for 5th graders, at the student level, interacting violence and students' observable characteristics.

Table 3.9 displays the results. Each column interacts separately violence and socioeconomic characteristics, while column 6 shows the more complete specification, where violence shocks are interacted simultaneously with different socioeconomic characteristics. The results that stand out in this last specification suggest that violence impacts relatively more the probability of dropping out of black and brown students. In Rio de Janeiro, nonwhite families and their children are more likely to live in slums and impoverished areas. The heterogeneous effect we find in Table 3.9 can reflect, for instance, that nonwhite students are more exposed to violent events - probably because their households, within slums, are closer to conflicts. On the other hand, race may capture other non-observable characteristics associated with poverty and socioeconomic disadvantages, which generally influence the individual's background and cognitive development. Thus, we can reasonably assume that nonwhite students are both those relatively more exposed to violence, and those who face more severe opportunities to learn and develop cognitive skills due to other disadvantages. This assumption suggests that violence might push the pool of remaining students at the end of the year towards a group of relatively high performing pupils. Therefore, selection bias might underestimate our estimated effects of violence exposure on student achievement and grade repetition. Thus, the effects of exposure to violence found in Table 3.7 might be a lower bound for the real impact.

3.6.3 What could explain violence effects on student achievement?

In this section we discuss two likely channels that can explain the negative effect of violence on student achievement. First, we explore the timing of the violence during the school year. Different effects of violence during classes or during vacations can shed light on the relative role of exposure to violence at home vis-a-vis exposure to violence at school. Thus, we can examine further to what extent violence effects work through the school vis-a-vis the household environment. Second, we study the relationship between violence and the patterns of teacher absenteeism and medical leaves. The relationship between violence and teachers' behavior during the year may help us to confirm whether violence work at the school environment through its human resources.

Violence Timing

We use our baseline model (3-14) in order to estimate the specifications shown in Table 3.10. The first column uses as variable of interest our measure of violence computing only the number of days with conflicts that occurred during school months (from February to June, and from August to October, up to Prova Brasil exams). In the second column, we consider only violent events that took place on January and July, months of school vacations. We see negative and significant effects in both specifications, with similar coefficients in terms of magnitude. This result suggests that the household environment channel might be of substantial importance, while we cannot reject the importance of the school environment channel. The result for the effects of violence during the vacation period suggest that exposure to violence can affect children outcomes via mental health and psychological disorders, as discussed in section 3.4.2.

Teacher absenteeism and medical leaves

In this section we examine the relationship between violence and teacher absenteeism.We use administrative records to build variables on teachers' unexcused absences and medical leaves at the school level. Unexcused absences are reported by the school's principal and is subject to endogeneity once the principal may under-report absenteeism in response to violence and safety threats. We believe that only events of long term absences are reported. Thus, unexcused absences may be interpreted as a combination of absences and turnover. Medical leaves are the main cause of teacher absenteeism in the municipal system. This type of leave is conceded only after medical examination supervised by the department of human resources of the municipal education authority.

For both types of teacher absenteeism, we calculate two variables at the school level for the years between 2004 and 2008. The first one is the sum of days of absence taken by all teachers of the school during the year. The second variable considers the average absence length, which is the former variable (sum of days of absence taken by all teachers of the school during the year) divided by the number of absence requests.[20]

These indicators are then used as dependent variables in our baseline model (3-14), which also includes as controls the number of teachers and students per school, as well as the the other controls used previously - school and year fixed effects and students' composition. Table 3.11 shows the results. Violence has a positive impact on unexcused absences but it does not affect medical leaves. Schools situated in the percentile 90 of the distribution of violence register more 3.06 days of absence during the year when compared to those schools situated in the percentile 10. This effect represents 17% of the sample average. The same variation in violence is associated with an increase in the average length of absences of 1.42 days, which represents 24% of the sample average. Taking into account that schools might under-report absences in response to external violent events, the evidence so far suggests that violence has disruptive effects on the school environment and student achievement through the human resources channel.

3.7 Robustness

In this section we present robustness checks for our results for achievement test scores in math presented in Table 3.7. Table 3.12 tests different empirical specifications. In the first two columns, we examine the role of past and future violent conflicts. In column 1 we include two lags of violence. Past violence coefficients help us to examine whether past violence is a significant input for contemporary achievement. We find no evidence that past violence is significantly associated with contemporaneous achievement, while the coefficient of contemporaneous violence remains at a similar magnitude and significant at 10%. Column 2 includes future violence in addition to lagged violence. This specification tests whether reverse causality is of concern and whether strict exogeneity assumption holds. We observe that coefficients of future shocks

[20]If a teacher absence for 30 days in the year divided in two spells, she enters twice in the denominator. If she absences 30 days uninterruptedly, she counts only once.

show no significant effects on contemporaneous achievement. However, significance of contemporaneous violence does not hold under this specification. This result must be driven by the fact that our sample is severely restricted only to variations between 2005 and 2007. Information from 2009 was discarded once we have no available data on violence for 2010 and 2011.

Specification in column 3 follows a weighted regression, using as weights the average number of 5th graders per school across time. This specification reduces the influence of small schools, in which few students contribute to the average score. We observe that our main result is robust to the use of weights. The point estimate of 0.97 is almost the same as the one presented in Table 3.7 (-1.09) and a little more noisier (standard-deviation of 0.52), but still significant at 10 percent confidence level.

Table 3.13 presents another set of robustness checks. We test whether our results are robust to sample selection, to the exclusion of outliers and to alternative measures of violence. In Panel A, we use the same measure of violence used throughout the paper. The first column shows our baseline result, the same displayed in Table 3.7, Panel A. In the second column we exclude outliers, i.e., the schools with extremely high records of violence. We define as outliers the schools from the top 1% of the respective violence distribution (the violence measure varies in each line). We see that the point estimate is marginally reduced, but still significant at 10%. In the third column, we exclude schools situated in Rio de Janeiro's Western Zone, where we are more subject to violence measurement error due to the presence of militia (see section 3.5.2). We observe that the point estimate almost double and significance is maintained at 5%. This result indicates that moving from percentile 10 to percentile 90 of violence distribution is associated with a 0.23 standard deviation decrease in math test scores. Specification in column 4 restricts the sample only to those schools located within 500 meters from a slum, while in column 5 we use only those schools distant at least 500 meters from a slum. We observe in column 4 that the significance and the magnitude of the coefficient are very similar to our baseline specification in the first column, while significance in column 5 vanishes. This result indicates that violence has very localized impacts.

In Panel B, we test whether another measure of violence, the standard deviation of our baseline measure of violence within years, is also associated with lower test scores. This measure is calculated as $ViolenceSD_{st} = \sum_j D_{sj} sd_{jt}$, where sd_{jt} is the standard deviation of the number of days with conflicts across months in each slum, and D_{sj} is the inverse of the distance between school s and slum j. As we can see from this specification, positive variations in this measure are associated with lower test scores no matter the sample we use. This result suggests that unevenly distributed violent events negatively affect student achievement.

In Panel C we use a much more flexible measure of violence in order to test not only robustness, but also the sensibility of our findings according to the distances between schools and slums. We use alternative measures of violence based on buffers of 5 meters, 250 meters and 500 meters around the school. Our three alternative measures of violence indicate the number of days with conflicts in slums within, respectively, a radius of 5, 250 and 500 meters from each school. In order to test the alternative specifications we use the baseline model (3-14), changing only samples and violence measures. We change the violence measure as indicated in the rows, and the sample as indicated in the columns of Table 3.13. We observe that in all columns the effect drops with buffer size, which is another result supporting the view that violence shocks have localized effects. The strongest effect is found for schools located within slums (buffers of 5 meters). Almost all measures are robust to alternative samples, particularly the 250-buffer. Row 4 in column A indicates that an additional day with conflict within 250 meters of distance from the school is associated with a 0.025 standard deviation decrease in math test scores. The results are robust to all measures of violence when we restrict our sample to those schools located within 500 meters from a slum.

Finally, in Panel D we alternatively use as a measure of violence the number of reports instead of the number of days with reports within the buffer of 250 meters. We find significant coefficients at 10%, with only column 3 as an exception.

Therefore, the set of results found so far seem robust to different measures of violence and sample selections. Given the magnitude of the estimated effects, the evidence suggests that exposure to local violence may have a substantial and pervasive role in children outcomes, particularly on those that study and live closer to the epicenters of violent events.

3.8 Conclusion

This paper investigates whether and how armed conflicts among drug gangs in Rio de Janeiro's slums affect children's educational outcomes. We explore time and geographical variation in localized violent events in order to identify causal effects of neighborhood violence on student achievement. By estimating a negative effect of violence on students' math scores, we find support to the view that exposure to violence has disruptive effects on children's outcomes .

This study provides several contributions. First, we develop a novel database, which contains precise information on whether and when drug conflicts happened, which allows a much better understanding of the problem. Second, we provide a better test to estimate violence impacts on schooling than the previous literature. This is possible once our dataset allows us to compare schools that are managed by the same municipal system and are, therefore, under the same rules and incentives, but vary in their distance to violence epicenters. In addition, the longitudinal structure of the data allows us to use school fixed effects and control for intrinsic characteristics of schools that are correlated with students and neighborhood demographics, leading to more precise estimates of violence impact.

Our results show that violence reduction should be a priority policy since its effects have a far-reaching impact beyond the great number of deaths caused by violent events. Our results support the view that violence accentuates the poverty trap, since it is particularly acute in poor areas. By decreasing the quality of learning in these areas, it makes the way out of poverty even harder for those children from disadvantaged households.

Table 3.1: Determinants of Drug Conflicts

	Dependent variable: Number of days with conflicts in the slum between 2004 and 2009								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Slum characteristics:									
steepness	0.27								0.15
	(0.03)***								(0.06)**
distance to main road		-1.03							-0.44
		(0.34)***							(0.17)**
area (1999)			0.04						0.03
			(0.00)***						(0.01)***
Neighborhood characteristics:									
population density				227.2					131.95
				(66.1)***					(52.18)**
population					-0.17				0.53
					(0.44)				(0.36)
% youngsters on						-65.0			-57.02
population (13-19 years)						(30.8)**			(45.35)
income pc							0.93		-1.90
							(1.14)		(1.37)
gini index								-3.64	8.10
								(6.43)	(5.77)
Observations	825	825	825	825	825	825	825	825	825
R^2	0.10	0.01	0.14	0.0	0.00	0.0	0.00	0.00	0.20

Notes: This Table presents coefficients from cross-section regressions of the number of days with conflicts between 2004 and 2009 on slum and neighborhood characteristics. The unit of analysis is the slum. Slum steepness is the standard deviation of the altitude within the slum calculated using GIS and NASA raster data. Distance to main road measures the smallest linear distance from slum border to a city main road, where the main roads are defined by Instituto Pereira Passos. Neighborhood information refers to population statistics, income per capita and gini index from the city neighborhood that the slum fall within. Asterisks indicate significantly different than zero at 99 (***), 95 (**), 90 (*) percent confidence.

Table 3.2: Disque-Denúncia Database Summary Statistics

Number of reports between 2004-2009		3,571		
Reporting gunfight		3,127		
on slums		2,884		92%
other places		243		8%
Number of slums		979		
with at least one report of gunfight		289		30%
without reports of gunfight		690		69%
Slums with conflicts				
	Number of reports		Number of days	
	per year	2004-2009	per year	2004-2009
mean	1.7	10	1.4	8
sd	5	18	3	14
p50	0	3	0	3
p90	4	26	4	22
max	85	146	41	96

Table 3.3: School Distribution

Proximity to slums	Number	%	% cumulative
Inside	25	3.4	3.4
Within 5-250 meters	321	43.6	47.0
Within 250-500 meters	192	26.1	73.1
More than 500 meters	198	26.9	100.0
Total	736		

Notes: This table presents the distribution of the 736 schools used in our sample by proximity to at least one slum. The proximity measure considers the linear distance between each school and the closest slum border.

Table 3.4: Education Summary Statistics

	Total		With violence =< 250 meters		Without violence =< 250 meters		
	Mean	sd	Mean	sd	Mean	sd	
Students' achievement and mobility:							
IDEB	4.65	(0.75)	4.36	(0.70)	4.75	(0.74)	***
math score	199.9	(18.9)	194.4	(17.6)	201.9	(19.0)	***
portuguese score	185.7	(16.9)	179.4	(15.4)	188.1	(16.8)	***
failure	0.09	(0.06)	0.1	(0.06)	0.09	(0.06)	***
dropout	0.04	(0.04)	0.05	(0.05)	0.03	(0.04)	***
school transfers	0.16	(0.06)	0.16	(0.06)	0.16	(0.06)	***
school admissions	0.18	(0.08)	0.16	(0.08)	0.19	(0.08)	***
Students' characteristics:							
% men	0.52	(0.03)	0.52	(0.03)	0.52	(0.03)	***
% white	0.37	(0.10)	0.33	(0.10)	0.38	(0.10)	***
mean age	9.25	(0.92)	9.11	(0.76)	9.3	(0.96)	***
% illiterate father	0.03	(0.02)	0.03	(0.03)	0.02	(0.02)	***
% illiterate mother	0.02	(0.02)	0.03	(0.02)	0.02	(0.02)	***
% mother and father at home	0.45	(0.13)	0.43	(0.12)	0.46	(0.13)	***
% live close to school	0.68	(0.24)	0.81	(0.16)	0.63	(0.24)	***
% evangelical	0.22	(0.08)	0.23	(0.08)	0.21	(0.08)	***
School characteristics:							
number students	537	(248)	574	(268)	524	(238)	***
pupil- teacher ratio	25.0	(6.7)	24.7	(6.5)	25.1	(6.7)	
% teachers with college degree	0.6	(0.24)	0.59	(0.22)	0.6	(0.24)	**
% principals < 4years on the job	0.04	(0.21)	0.05	(0.22)	0.04	(0.20)	
% with science lab	0.1	(0.30)	0.07	(0.25)	0.11	(0.32)	***
% with computer lab	0.38	(0.49)	0.45	(0.50)	0.36	(0.48)	**
% with free meal	0.99	(0.10)	0.99	(0.11)	0.99	(0.09)	
% with sport court	0.59	(0.49)	0.6	(0.49)	0.59	(0.49)	
Principal reported problem with:							
suspended classes	0.05	(0.21)	0.14	(0.35)	0.01	(0.11)	***
teachers' absence	0.04	(0.20)	0.04	(0.20)	0.04	(0.20)	
students' absence	0.07	(0.26)	0.10	(0.30)	0.06	(0.24)	*
drug consumption close to school	0.24	(0.43)	0.39	(0.49)	0.18	(0.39)	***
drug traffic close to school	0.22	(0.42)	0.44	(0.50)	0.14	(0.34)	***

Notes: This table presents a comparison of the average students' achievement and mobility, students' characteristics, school characteristics and principals' reports between schools with and without violence within 250 meters from the school. Column 1 reports the means for the 736 schools used in our sample. Column 3 reports the means for the 199 schools exposed to violence within 250 meters in any moment between 2004 and 2009, while column 5 reports the means for the other 537 schools non-exposed to violence in the period. Columns 2,4 and 6 report standard errors for each sample. Asterisks presented in column 7 indicates whether the difference in means between columns 3 and 5 are significantly different than zero at 99 (***), 95 (**), 90 (*) percent confidence.

Table 3.5: Testing for under-reporting

AISP		2004	2005	2006	2007	AISP		2004	2005	2006	2007
1	Homicide rate	118	83	71	74	16	Homicide rate	129	149	150	170
	Pred homicide	184	266	205	174		Pred homicide	217	165	180	183
	Under-reporting	0	0	0	0		Under-reporting	0	0	0	0
2	Homicide rate	36	20	33	23	17	Homicide rate	80	49	59	38
	Pred homicide	97	100	66	58		Pred homicide	151	150	107	54
	Under-reporting	0	0	0	0		Under-reporting	0	0	0	0
3	Homicide rate	153	135	166	199	18	Homicide rate	138	150	133	123
	Pred homicide	221	215	252	322		Pred homicide	88	105	66	72
	Under-reporting	0	0	0	0		Under-reporting	1	1	1	1
4	Homicide rate	45	33	43	22	19	Homicide rate	16	19	11	12
	Pred homicide	82	70	60	63		Pred homicide	91	77	86	72
	Under-reporting	0	0	0	0		Under-reporting	0	0	0	0
5	Homicide rate	38	55	42	37	22	Homicide rate	209	137	110	115
	Pred homicide	76	60	55	49		Pred homicide	140	80	81	91
	Under-reporting	0	0	0	0		Under-reporting	1	1	1	1
6	Homicide rate	54	67	79	88	23	Homicide rate	37	41	33	28
	Pred homicide	186	155	143	169		Pred homicide	101	140	91	63
	Under-reporting	0	0	0	0		Under-reporting	0	0	0	0
9	Homicide rate	617	532	480	454	27	Homicide rate	238	182	232	231
	Pred homicide	178	205	273	345		Pred homicide	155	80	164	151
	Under-reporting	1	1	1	1		Under-reporting	1	1	1	1
13	Homicide rate	17	14	14	18	31	Homicide rate	50	46	51	38
	Pred homicide	68	57	45	49		Pred homicide	68	57	45	49
	Under-reporting	0	0	0	0		Under-reporting	0	0	1	0
14	Homicide rate	372	368	414	339	39	Homicide rate	305	326	344	327
	Pred homicide	412	301	444	294		Pred homicide	136	123	102	77
	Under-reporting	0	1	0	1		Under-reporting	1	1	1	1

Notes: This Table presents the actual and predicted homicide rate of each Área Integrada de Segurança Pública (AISP), which is a division of Rio de Janeiro used by Police Authority to provide crime statistics. In order to calculate predicted homicide, we run yearly regressions of homicide rate on the number of days with reports about armed conflicts. We then used the estimated coefficient to generate predicted homicide. Under-reporting indicates whether the predicted homicide rate is lower than the actual homicide rate.

Table 3.6: Principals' reports about school violence

Caused by:	Teachers robbed inside school non-students	Teachers robbed inside school students	Drug consumption outside school non-students	Drug traffic inside school non-students	Drug traffic outside school non-students	Gangs outside school	Gangs inside school
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Violence	-0.013	0.001	0.058	-0.003	0.080	0.050	0.010
	(0.007)*	(0.007)	(0.026)**	(0.009)	(0.026)***	(0.018)***	(0.010)
Observations	570	418	595	564	593	650	646
R^2	0.049	0.080	0.088	0.035	0.136	0.088	0.040

Notes: This table reports the results of cross-section regressions of the dependent variable indicated in each column on DD violence measure (the logarithm of the weighted sum of all slums' days with armed conflicts, where the weight is the inverse of the distance from each slum to the school). The dependent variables are binary indicators for whether the cited event happened in the school in 2007. We include as controls principals' characteristics (how long she is on the job in the school, previous experience as principal, age range, education, whether she has another job), students' average characteristics (share of whites, average asset index, share that live close to school) and school characteristics (number of students and whether the school has kitchen, principal room, science lab, computer lab, free meals, sport court, and teachers' room). Robust standard errors in parentheses, significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 3.7: Achievement Regressions at the School Level - Math and Language Test Scores

	Effects of violence on Math and Language test scores		
	(1)	(2)	(3)
	Panel A: Math		
Violence	-1.025	-0.261	-1.094
	(0.381)***	(0.352)	(0.533)**
	Panel B: Language		
Violence	-0.662	-0.127	-0.272
	(0.380)*	(0.328)	(0.495)
Common Specifications:			
Observations	2,125	2,117	2,117
Number of schools	736	736	736
School FE	NO	NO	YES
Students composition	NO	YES	YES
Year FE	YES	YES	YES

Notes: Robust standard errors in parentheses, adjusted for clustering at the school level in all specifications. Significance: *** p<0.01, ** p<0.05, * p<0.1. Dependent variables are school's average achievement test scores in math (Panel A) and language (Panel B) for 5th graders in years 2005, 2007 and 2009. All regressions include year fixed effects. Columns (2) and (3) include number of 5th graders per school and controls for 5th graders' socioeconomic composition, averaged at the school level: average age, share of white students, share of boys, students' mothers education (share of mothers with incomplete primary, complete primary, complete secondary and college). Only column (3) includes school fixed effects. Variable of interest (violence) is the logarithmic of the sum of days of conflicts per slum weighed by the distance between school and slums.

Table 3.8: Grade Repetition, Dropout and Mobility in the 5th Grade

	Grade Repetition	Dropout	Transfers (out of school)	Admissions (in)
	(1)	(2)	(3)	(4)
Violence	1.253	0.188	-0.099	-0.900
	(0.605)**	(0.087)**	(0.278)	(0.624)
Observations	4,215	4,218	4,218	4,218
Number of schools	736	736	736	736

Notes: Robust standard errors in parentheses clustered at the school level. Significance: *** p<0.01, ** p<0.05, * p<0.1. Dependent variables are schools' average rates of grade repetition (share of students that repeated the grade over total students at the end of the year in that grade), dropout (share of dropout over total students at the beginning of the year), transfers and new admissions (both analogously to dropout) among 5th graders. The sample includes years between 2004 and 2009. All regressions include school and year fixed effects, and students' composition per grade and year: number of students, average age, share of white students, share of boys, students' mothers education (share of mothers with incomplete primary, complete primary, complete secondary and college). Variable of interest (violence) is the logarithmic of the sum of days of conflicts per slum weighed by the distance between school and slums.

Table 3.9: Heterogeneity in Students' Dropout

	Dependent variable: Dropout in 5th grade					
	(1)	(2)	(3)	(4)	(5)	(6)
Violence (V)	0.0006	0.0004	0.0008	-0.0010	0.0142	0.0125
	(0.0010)	(0.0010)	(0.0011)	(0.0012)	(0.0100)	(0.0100)
V * (High educated mother)		0.0018				0.0013
		(0.0011)*				(0.0009)
V * Boys			-0.0003			-0.0001
			(0.0010)			(0.0009)
V * (Brown or Black)				0.0022		0.0026
				(0.0009)**		(0.0009)***
V * (Age)					-0.0012	-0.0013
					(0.0009)	(0.0009)
Observations	392,843	392,843	392,843	392,843	392,843	392,843

Notes: This table presents coefficients from regressions of student's dropout in the 5th grade on violence and violence interacted with students' characteristics. Observations are at the student level. Variable of interest (violence) is the logarithmic of the sum of days of conflicts per slum weighed by the distance between school and slums. All regressions include school and year FE and controls for number of students at 5th grade and students' characteristics: age, dummy indicating boys, race (brown or black) dummy, high educated mother (high school or college degree) dummy. Sample covers years between 2004 and 2009. Standard errors adjusted for clustering at the school level in parentheses, significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 3.10: Violence Timing

	Dependent variable: Math test scores	
	(1)	(2)
Violence during classes	-1.177	
	(0.546)**	
Violence during vacations		-1.278
		(0.568)**
Observations	2117	2117
Number of schools	736	736

Notes: This table presents coefficients from regressions of Prova Brasil math score on violence. The variable violence is the logarithmic of the sum of days of conflicts per slum weighed by the distance between school and slums. Violence during school period is the number of days with conflicts from February to June and from August to October (until Prova Brasil application). Violence during vacation period includes the number of days with conflicts in January and July. All regressions include school and year fixed effects and controls for the number of students at fifth grade and fifth graders' average characteristics (share of men, average age, share of whites, mother's education). The period of analysis covers the years in which Prova Brasil was applied (2005, 2007 and 2009). Standard errors adjusted for clustering at the school level in parentheses, significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 3.11: Teachers Absenteeism

	Unexcused Absences		Medical Leaves	
	Number of Days of Absence	Absence Length	Number of Days of Absence	Absence Length
	(1)	(2)	(3)	(4)
Violence	1.900	0.856	3.170	0.440
	(0.822)**	(0.423)**	(8.609)	(0.742)
Y mean	18.18	6.03	394.6	28.4
Observations	4,035	4,035	4,035	4,035
Number of schools	736	736	736	736

Notes: This table presents coefficients from regressions of teachers' absenteeism on violence. Dependent variables are measured at the school level as follows. Number of days of absence refer to the sum of days of absence taken by all teachers of the school during the year. Column (1) refers to unexcused absences while column (3) refers to absences due to health problems. Analogously, in columns (2) and (4) absence length refers to the sum of days of absence taken by all teachers of the school during the year divided by the number of absence requests. Violence is the logarithmic of the sum of days with conflicts in each slum weighed by the distance between school and slums. All regressions include school and year fixed effects, and controls for the number of teachers in the school, the number of students at fifth grade and fifth graders' average characteristics (share of men, average age, share of whites, mother's education). Sample covers years between 2004 and 2009. Standard errors adjusted for clustering at the school level in parentheses, significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3.12: Econometric Specification Robustness Checks

	Dependent variable: Math test scores		
	(1)	(2)	(3)
Violence t-2	0.383	0.518	
	(0.641)	(0.930)	
Violence t-1	0.193	0.051	
	(0.771)	(1.234)	
Violence t	-1.024	-0.476	-0.982
	(0.527)*	(0.886)	(0.511)*
Violence t+1		-0.650	
		(1.272)	
Violence t+2		0.381	
		(0.977)	
Observations	2117	1411	2117
Number of schools	736	736	736

Notes: This table presents coefficients from regressions of Prova Brasil math score on violence. The variable violence is the logarithmic of the sum of days of conflicts per slum weighed by the distance between school and slums. All regressions include school and year fixed effects and controls for the number of students at fifth grade and fifth graders' average characteristics (share of men, average age, share of whites, mother's education). In column 3, we weighted the regression by the average number of 5th graders over the sample period. The period of analysis cover the years in which Prova Brasil was applied (2005, 2007 and 2009), except by column 5 which include only 2005 and 2007. Standard errors adjusted for clustering at the school level in parentheses, significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 3.13: Sample and Measure of Violence Robustness Checks

Dependent variable:	Math test scores				
Sample:	Full Sample	Exclude Outliers	Without Western Zone	Slum Distance < 500m	Slum Distance > 500m
	(1)	(2)	(3)	(4)	(5)
			Panel A: Baseline measure of violence		
Violence	-1.094	-1.011	-1.922	-1.075	1.660
(baseline)	(0.533)**	(0.591)*	(0.774)**	(0.537)**	(4.908)
Schools	736	729	418	538	198
			Panel B: Violence standard deviation		
Violence	-0.004	-0.009	-0.003	-0.004	-0.029
Sd deviation	(0.001)***	(0.004)**	(0.001)**	(0.001)***	(0.537)
Schools	736	729	418	538	198
			Panel C: Buffers (number of days)		
Buffer 5 meters	-0.413	-0.457	-0.355	-0.424	
(num of days)	(0.192)**	(0.688)	(0.204)*	(0.190)**	
Schools	736	728	418	538	
Buffer 250 meters	-0.334	-0.344	-0.264	-0.339	
(num of days)	(0.145)**	(0.172)**	(0.157)*	(0.145)**	
Schools	736	728	418	538	
Buffer 500 meters	-0.140	-0.154	-0.163	-0.150	
(num of days)	(0.083)*	(0.102)	(0.093)*	(0.082)*	
Schools	736	728	418	538	
			Panel D: Buffers (number of reports)		
Buffer 250 meters	-0.222	-0.231	-0.176	-0.224	
(num of reports)	(0.119)*	(0.131)*	(0.129)	(0.118)*	
Schools	736	729	418	538	

Notes: Each table entry represents a regression coefficient of the math test scores on violence. The violence measure is indicated in each line, while in the columns we vary the sample used. The baseline measure (Panel A) is the logarithmic of the sum of days of conflicts per slum weighed by the distance between school and slums. Violence standard deviation (Panel B) is the variation across the number of days with conflicts in each month of the year. The variables buffers (Panel C) indicate the number of days with conflicts in slums within, respectively, 5 meters, 250 meters and 500 meters from school. In Panel D buffer 250m refers to the number of reports about conflicts in slums within 250 meters from school. All regressions include school and year fixed effects and controls for the number of students at 5th grade and 5th graders' average characteristics (share of boys, average age, share of whites, mother's education). Column 2 excludes outliers, the top 1% of the distribution of the respective measure of violence. Column 3 excludes slums and schools from Rio de Janeiro's Western Zone. The period of analysis cover the years in which Prova Brasil was applied (2005, 2007 and 2009). Standard errors adjusted for clustering at the school level in parentheses, significance: *** p<0.01, ** p<0.05, * p<0.1.

Table 3.14: Grade Repetition, Dropout and Mobility - 1st to 4th grades

	Effects of Violence per Grade Rates Grade Repetition, Dropout and Mobility			
	1st	2nd	3rd	4th
	Panel A: Grade Repetition			
Violence	0.020	-0.034	0.151	-0.027
	(0.146)	(0.091)	(0.310)	(0.243)
Observations	4,069	4,085	4,115	4,280
Number of schools	700	703	701	732
	Panel B: Dropout			
Violence	0.195	0.025	0.159	0.041
	(0.157)	(0.095)	(0.080)**	(0.097)
Observations	4,069	4,086	4,116	4,281
Number of schools	700	703	701	732
	Panel C: Transfers (out)			
Violence	0.088	0.127	0.132	0.105
	(0.281)	(0.197)	(0.231)	(0.283)
Observations	4,069	4,086	4,116	4,281
Number of schools	700	703	701	732
	Panel D: New Admissions (in)			
Violence	0.413	0.113	0.126	0.416
	(0.502)	(0.273)	(0.324)	(0.369)
Observations	4,069	4,086	4,116	4,281
Number of schools	700	703	701	732

Notes: Robust standard errors in parentheses clustered at the school level. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Dependent variables are schools' rates of grade repetition (share of students that repeated grade over total students at the end of the year), dropout (share of dropout over total students at the beginning of the year), transfers and new admissions (both analogously to dropout). All dependent variables are per grade and the sample includes years between 2004 and 2009. All regressions include school and year fixed effects, and students' composition per grade and year: number of students, average age, share of white students, share of boys, students' mothers education (share of mothers with incomplete primary, complete primary, complete secondary and college). Variable of interest (violence) is the logarithmic of the sum of days of conflicts per slum weighed by the distance between school and slums. Samples change across columns because some schools do not offer all grades every year.

Figure 3.1: Slum and School Distribution

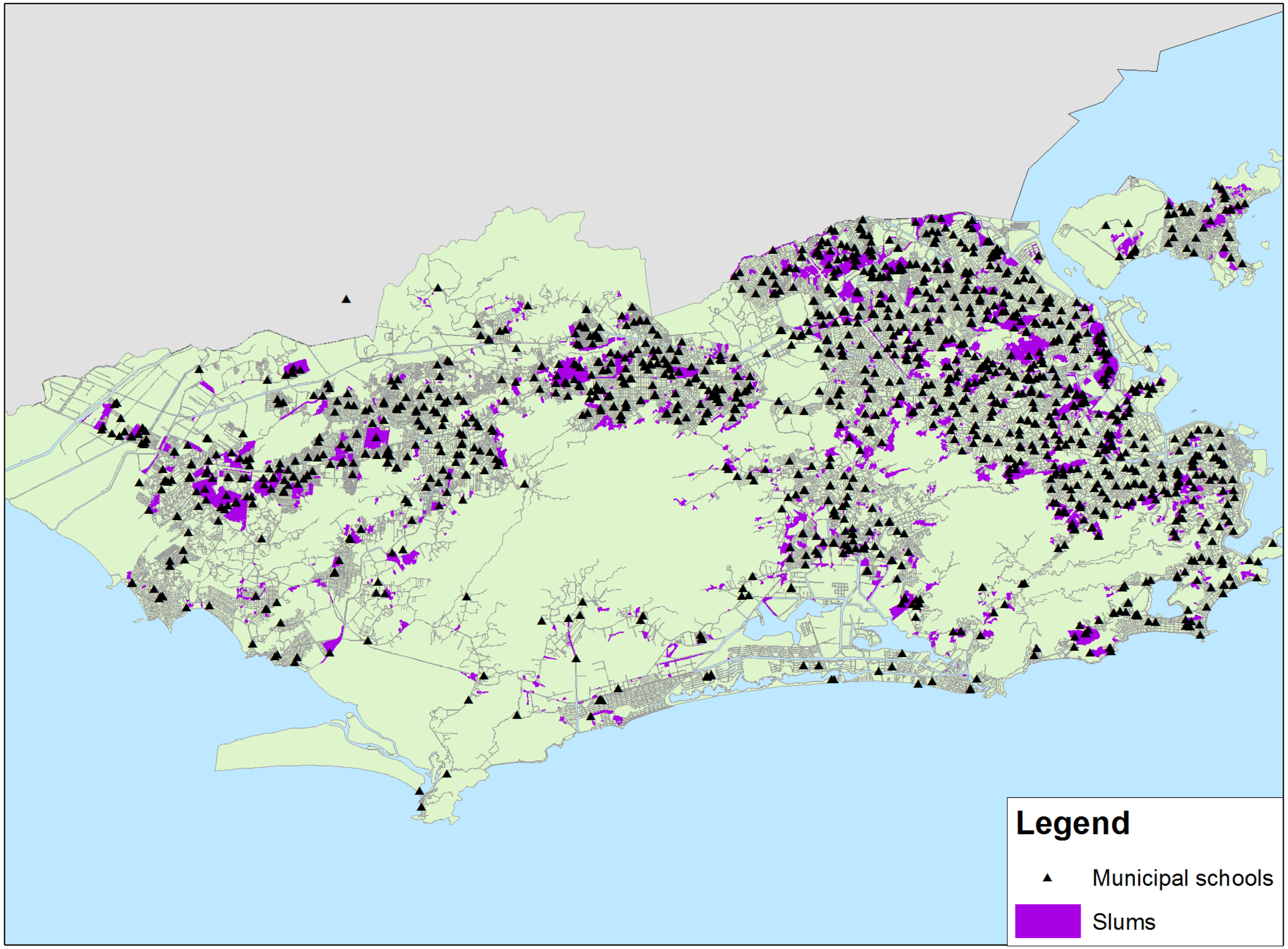


Figure 3.2: Number of Days with Reports about Gunfight 2004-2009

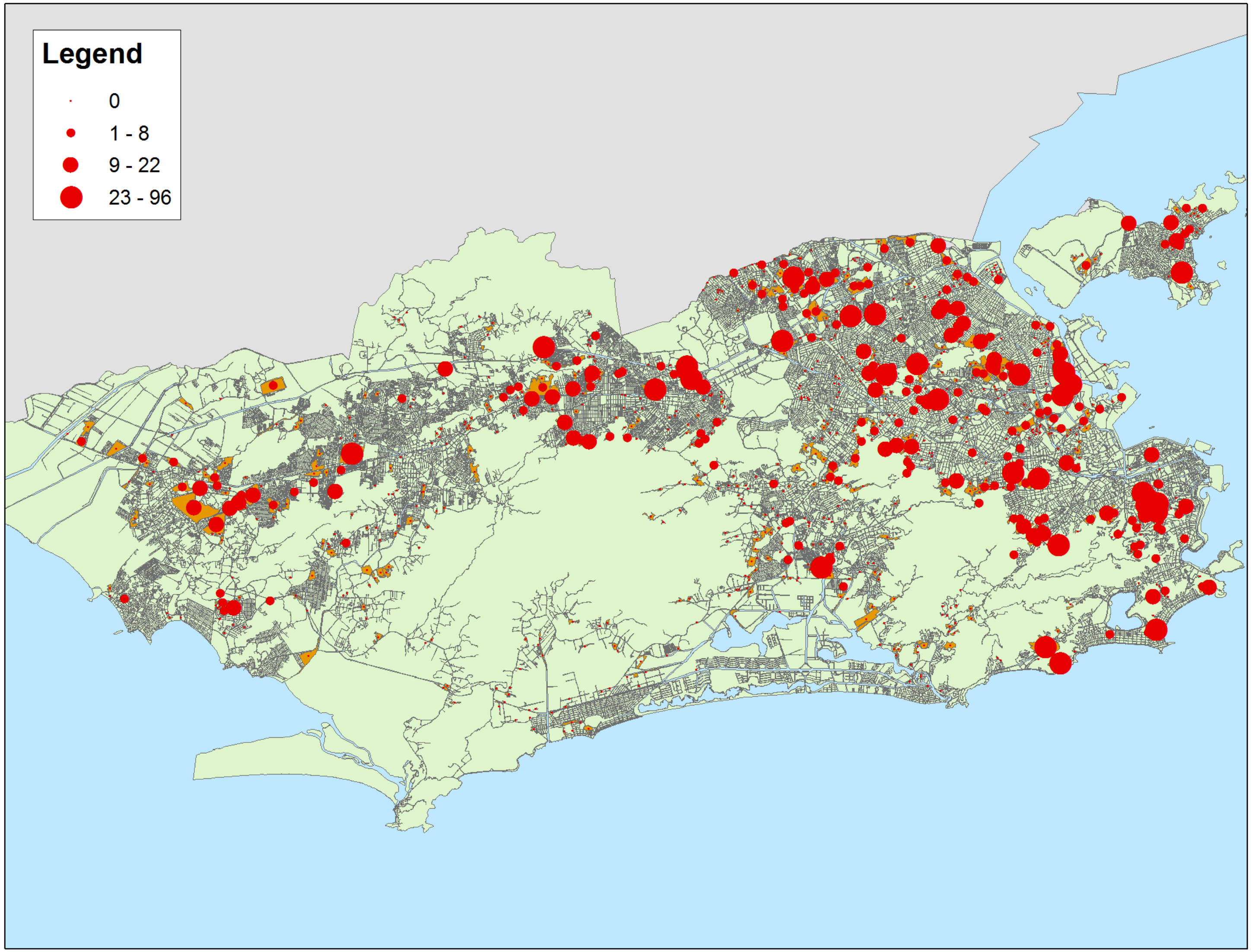
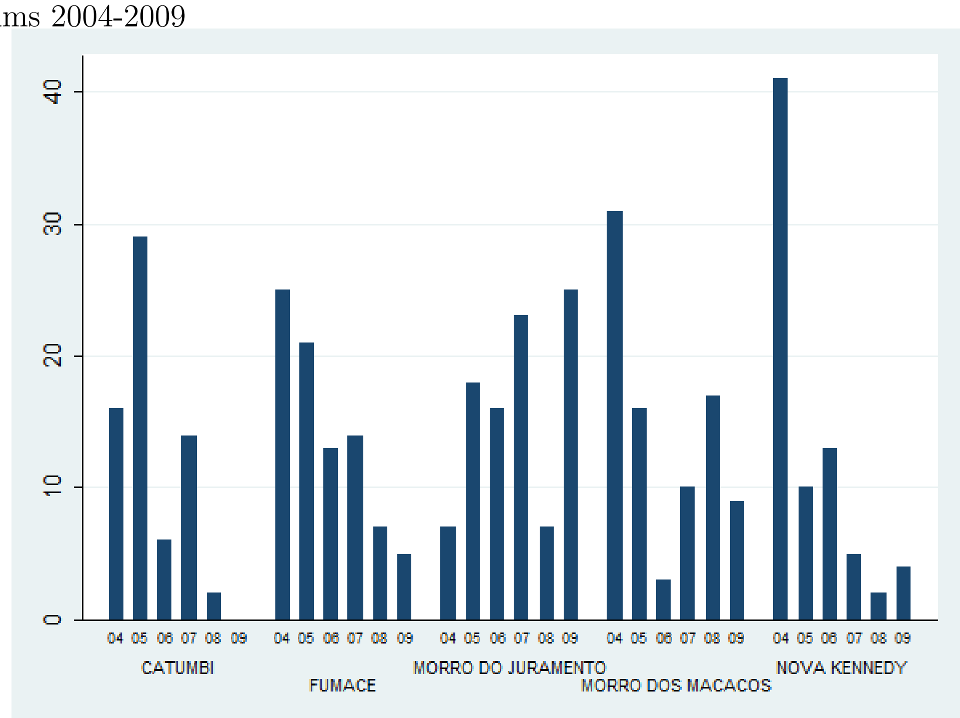


Figure 3.3: Number of Days with Reports about Gunfight per Year in Selected Slums 2004-2009



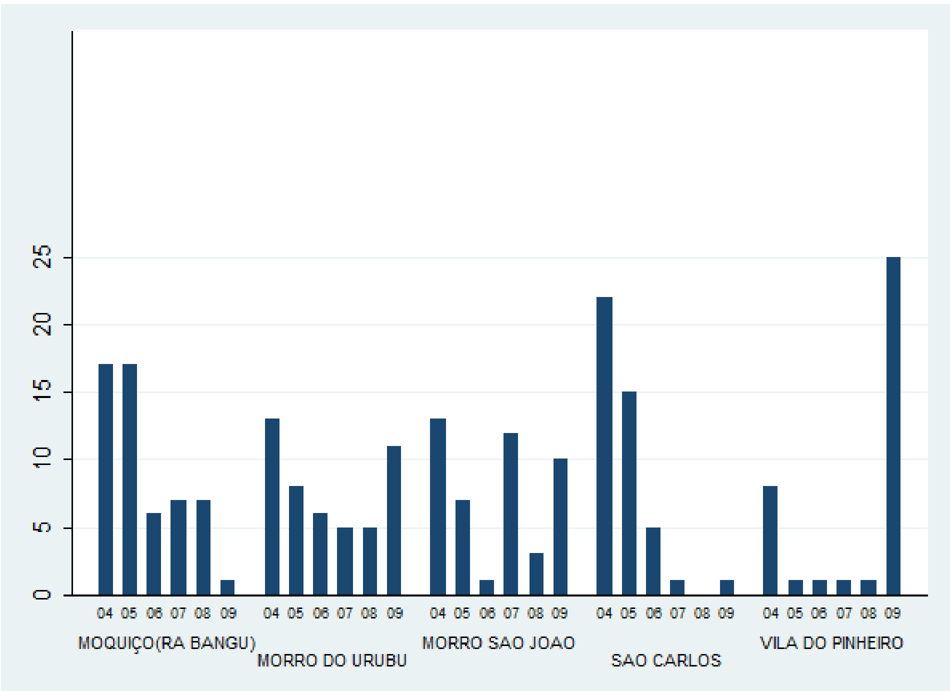
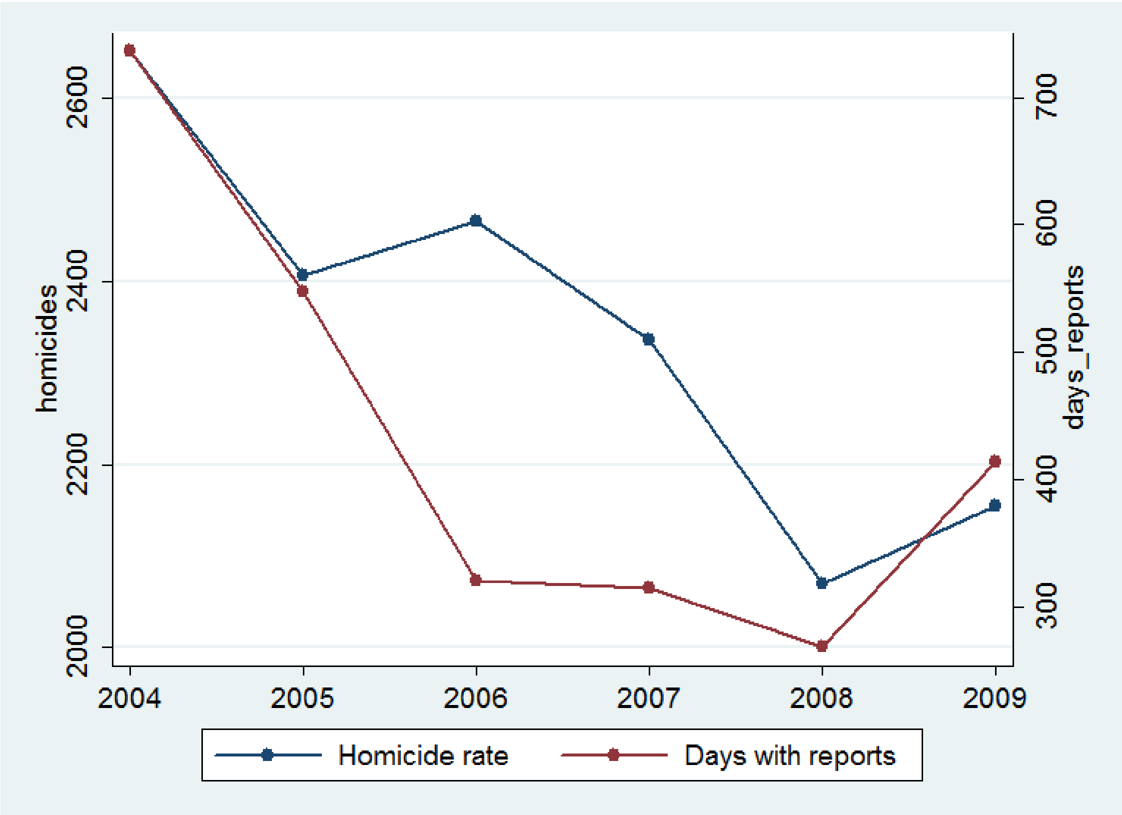
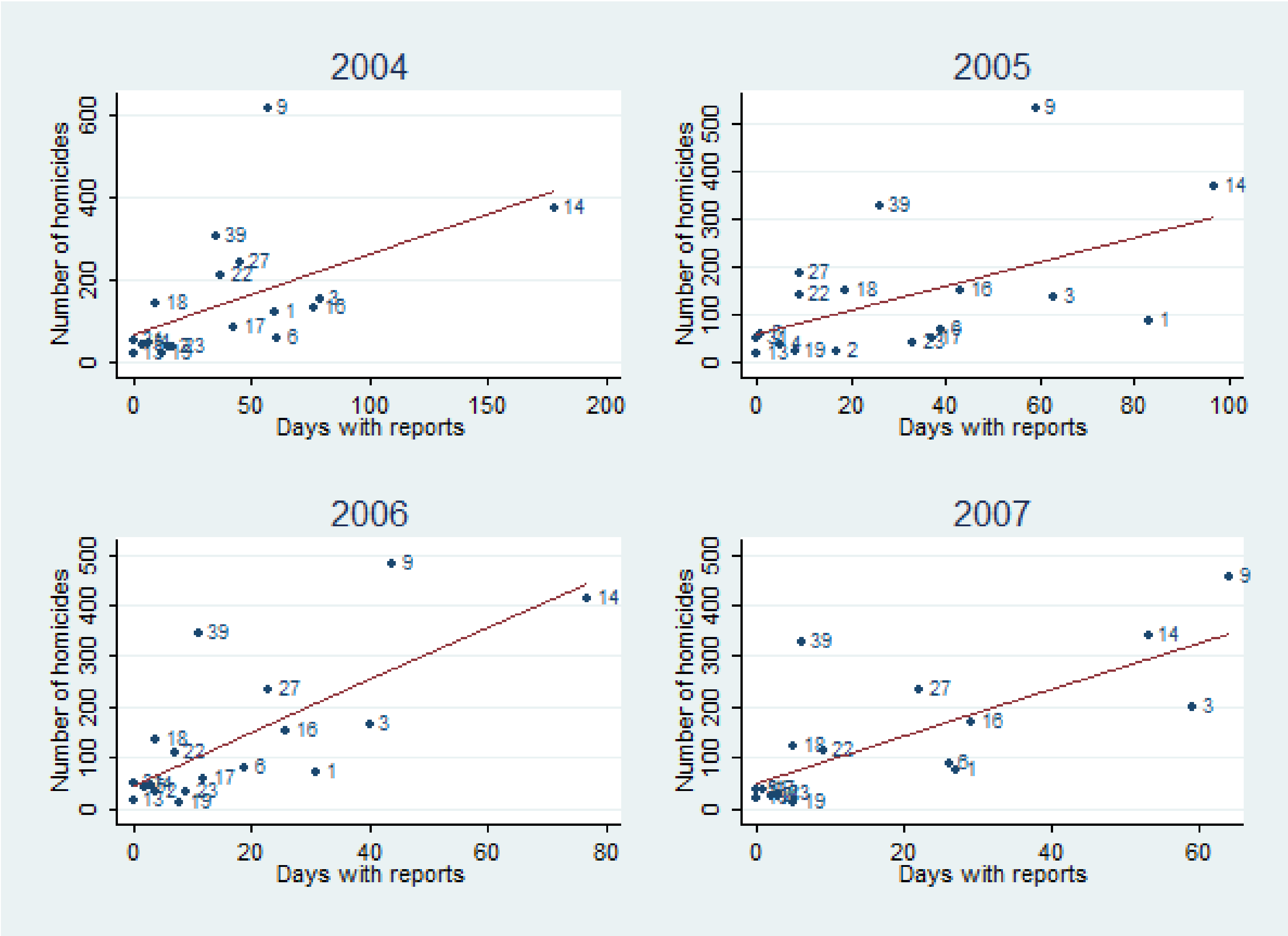


Figure 3.4: Homicides and Number of Days with Conflicts 2004-2009



Notes: This figure compares the number of homicides and the levels of violence documented in Disque-Denúncia reports between 2004 and 2009. The left y-axis indicates the number of homicides in the city of Rio de Janeiro. The right y-axis indicates the sum of the number of days with reports about gunfight in all Rio de Janeiro's slums.

Figure 3.5: Homicides and Number of Days with Conflicts per AISP



Notes: This figure shows the correlation between the number of homicides in the city of Rio de Janeiro and the number of days with conflicts in Rio de Janeiro's slums. Both measures are aggregated per AISP (the city division used by the police department). Each panel indicates a different year.