This document helps to replicate the analysis in *Drug Battles and School Achievement: Evidence from Rio de Janeiro's Favelas*. We provide below a description of each database and its variables, along with guidelines to import them using diferent softwares. We also indicate which code files should be used to replicate tables and figures.

1 Databases

1.1 School's database

File: data_schools.txt

Import: the database is comma separated. The first row contains the names of the variables. The imported data must have 2,352 observations and 50 variables.

Description: the file contains information (2003-2009) about 336 elementary schools of the municipal system that participated in at least two Prova Brasil editions between 2005 and 2009, and are located within 250 meters from a favela in Rio de Janeiro. Table 1 presents the labels to all variables in the database.

Data source: Instituto Nacional de Estudos e Pesquisas Educacionais (INEP)

1.2 Student's database

File: data_students.txt

Import: the database is comma separated. The first row contains the names of the variables. The imported data must have 163,293 observations and 116 variables.

Description: the file contains information (2005,2007,2009) about the fifth graders students from 736 schools of the municipal system that participated in at least two editions of Prova Brasil between 2005 and 2009. Table 2 presents the labels to all variables in the database.

Data source: Instituto Nacional de Estudos e Pesquisas Educacionais (INEP)

1.3 Teacher's database

File: data_teachers.txt

Import: the database is comma separated. The first row contains the names of the variables. The imported data must have 20,110 observations and 14 variables.

Description: the file contains information (2007-2009) about teachers of the 336 elementary schools of the municipal system that participated in at least two Prova Brasil editions between 2005 and 2009, and are located within 250 meters from a favela in Rio de Janeiro. Table 3 presents the labels to all variables in the database.

Data source: Instituto Nacional de Estudos e Pesquisas Educacionais (INEP)

1.4 Student's mobility database

File: data_movements.txt

Import: the database is comma separated. The first row contains the names of the variables. The imported data must have 1,136,594 observations and 26 variables.

Description: the file contains information (2003-2009) about student mobility in the 336 elementary schools of the municipal system that participated in at least two Prova Brasil editions between 2005 and 2009, and are located within 250 meters from a favela in Rio de Janeiro. Table 4 presents the labels to all variables in the database.

Data Source: Rio de Janeiro's Secretaria Municipal de Educação (SME)

1.5 Report's databases

File: data_reports_favela.txt

Import: the database is comma separated. The first row contains the names of the variables. The imported data must have 351,140 observations and 7 variables.

Description: the file contains the number of anonymous calls to Disque-Denuncia reporting gunfights between drug gangs per favela and day (2003-2009), as well as the number of calls

that indeed report gunfight between drug gangs based on the reading of all reports. Users should have in mind that *id_favela=8888* indicate locations that were not associated with a specific favela. Table 5 presents the labels to all variables in the database.

Data source: Disque-Denúncia (DD)

File: data_reports_aisp.txt

Import: the database is comma separated. The first row contains the names of the variables. The imported data must have 126 observations and 6 variables.

Description: the file contains the number of anonymous calls to Disque-Denuncia reporting gunfights between drug gangs per AISP and year (2003-2009), as well as the number of homicides per AISP and year. Table 6 presents the labels to all variables in the database.

Data source: Disque-Denúncia (DD)

Instituto de Segurança Pública do Rio de Janeiro (ISP)

2 Analysis

File: monteiro_rocha_replicate.do

Description: this do-file provides the code to generate all the tables and figures presented in the main paper. Users should create a folder called "Replication" which holds three further subfolders: "Databases", "Tables" and "Figures". Place all the five databases described in section 1 in the subfolder "Databases". Replace the text between quotation marks on the first line of the do-file with the directory path to the "Replication" folder.

Files: monteiro_rocha_replicate_conley.do

ols_spatial_HAC.ado

Description: the do-file provides the code to generate the conley standards errors presented in tables at the main paper and at the web-appendix. Users should create a forth subfolder

called “Conley” in the “Replication” folder and replace the text between quotation marks on the first line of the do-file with the directory path to the “Replication” folder. Additionally, users should place the file ols_spatial_HAC.ado in Stata’s root directory (/ado/base/o).

3 Software

Stata/SE 12.0

Table 1: School's database dictionary

id_school	**school identification number**
year	**year of Prova Brasil editions**
absences_average	total number of absences per number of teachers
age_teachers	teacher's average age
age_test	average age of students who took the test
computer_lab	1 = school has a computer lab
count	2 = school participated in both 2007 and 2009 editions of Prova Brasil
d_violence_250m_classes	violence during school period (Mar-Nov) within 250 meters
d2	indicator for >=2 days of violence (march-november)(250m)
d2Ncontiguous	indicator for >=2 day non-contiguous violence (march-november)
day_contiguous	number of days of contiguous violence (14 day window)
dcontiguous	indicator for >=2 day contiguous violence (march-november)
dropped_test	number of students that have dropped out of school in the past per number of students who took the test
grad_teachers	share of teachers with graduate degree
higheduc_mother_test	number of students with highly educated mother, who took the test, per number of students who took the test
interrup	class interruption reported as a problem (yes=1 no=0 missing=.)
interrup_miss	class interruption reported as a problem (missing=1 yes&no=0)
interrup_yes	class interruption reported as a problem (yes=1 no&missing=0)
kitchen	1 = school has a kitchen

Table 1: School's database dictionary (continuing)

male_teachers	share of male teachers
males_test	number of male students who took the test per number of students who took the test
new_principal	principal for less than 2 years (yes=1 no=0 missing=.)
new_principal_miss	principal for less than 2 years (missing=1 yes&no=0)
new_principal_yes	principal for less than 2 years (yes=1 no&missing=0)
number_teachers	number of teachers in the school
number_stud_school	number of students in the school
onleave_average	total number of days on medical leave per number of teachers
POINT_X	school latitude
POINT_Y	school longitude
principal_office	1 = school has principal's office
repeated_test	number of students that have failed a grade in the past per number of students that took the test
school_lunch	1 = school offers lunch
science_lab	1 = school has a science lab
share_absent	share of absent teachers
share_onleave	share of teachers on medical leave
student_absence	student absence reported as a problem (yes=1 no=0 missing=.)
student_absence_miss	student absence reported as a problem (missing=1 yes&no=0)
student_absence_yes	student absence reported as a problem (yes=1 no&missing=0)
teacher_office	1 = school has teacher's office

Table 1: School's database dictionary (continuing)

threat_student	threat to the life of students (by external agents) reported as a problem (yes=1 no=0 missing=.)
threat_student_miss	threat to the life of students (by external agents) reported as problem (missing=1 yes&no=0)
threat_student_yes	threat to the life of students (external agents) reported as a problem (yes=1 no&missing=0)
threat_teacher	threat to the life of teachers or staff (by external agents) reported as a problem (yes=1 no=0 missing=.)
threat_teacher_miss	threat to the life of teachers or staff (by external agents) reported as a problem (missing=1 yes&no=0)
threat_teacher_yes	threat to the life of teachers or staff (by external agents) reported as a problem (yes=1 no&missing=0)
turnover	teachers turnover reported as a problem (yes=1 no=0 missing=.)
turnover_miss	teachers turnover reported as a problem (missing=1 yes&no=0)
turnover_yes	teachers turnover reported as a problem (yes=1 no&missing=0)
undergrad_teachers	share of teachers with undergraduate diploma
white_test	number of white students who took the test per number of students who took the test

Table 2: Student's database dictionary

id_school	**school identification number**
year	**year of Prova Brasil editions**
id_class	**class identification number**
id_student	**student identification number**
age	student age (missing = -1)
age_11	child is 11 years old or younger
age_class	average age of students in the class
age_incorrect	child is 12 years old or older
age_m	student age (missing = .)
boy	student is a boy (missing = -1)
boy_m	student is a boy (missing = .)
boys_class	share of boys in the class
computer_lab	1 = school has a computer lab
d_violence_100m_classes	violence during school period (Mar-Nov) within 100 meters
d_violence_150m_classes	violence during school period (Mar-Nov) within 150 meters
d_violence_200m_classes	violence during school period (Mar-Nov) within 200 meters
d_violence_250m_classes	violence during school period (Mar-Nov) within 250 meters
d_violence_300m_classes	violence during school period (Mar-Nov) within 300 meters
d_violence_350m_classes	violence during school period (Mar-Nov) within 350 meters
d_violence_400m_classes	violence during school period (Mar-Nov) within 400 meters
d_violence_450m_classes	violence during school period (Mar-Nov) within 450 meters

Table 2: Student's database dictionary (continuing)

d_violence_500m_classes	violence during school period (Mar-Nov) within 500 meters
d_violence_50m_classes	violence during school period (Mar-Nov) within 50 meters
d_violence_750m_classes	violence during school period (Mar-Nov) within 750 meters
d1	indicator for 1>= days of violence (march-november)(250m)
d100m	indicator for slum border closer than 100m
d150m	indicator for slum border closer than 150m
d2	indicator for 2>= days of violence (march-november)(250m)
d2_100m	indicator >=2 days Mar-Nov within 100 meters)
d2_150m	indicator >=2 days Mar-Nov within 150 meters)
d2_1trim	indicator >=2 days of violence btw December and February
d2_200m	indicator >=2 days Mar-Nov within 200 meters)
d2_250m	indicator >=2 days Mar-Nov within 250 meters)
d2_2trim	indicator >=2 days of violence btw March and May
d2_300m	indicator >=2 days Mar-Nov within 300 meters)
d2_350m	indicator >=2 days Mar-Nov within 350 meters)
d2_3trim	indicator >=2 days of violence btw June and August
d2_400m	indicator >=2 days Mar-Nov within 400 meters)
d2_450m	indicator >=2 days Mar-Nov within 450 meters)
d2_4trim	indicator >=2 days of violence btw September and November
d2_500m	indicator >=2 days Mar-Nov within 500 meters)
d2_50m	indicator >=2 days Mar-Nov within 50 meters)
d2_5m	indicator >=2 days Mar-Nov within 5 meters)

Table 2: Student's database dictionary (continuing)

d2_750m	indicator >=2 days Mar-Nov within 750 meters)
d2_fall	indicator >=2 days fall)
d2_spring	indicator >=2 days spring)
d2_vacation	indicator >=2 days vacation)
d200m	indicator for slum border closer than 200m
d250m	indicator for slum border closer than 250m
d2Ncontiguous	indicator for >=2 day non-contiguous violence Mar-Nov
d3	indicator for 3>= days of violence (march-november)(250m)
d300m	indicator for slum border closer than 300m
d350m	indicator for slum border closer than 350m
d4	indicator for 4>= days of violence (march-november)(250m)
d400m	indicator for slum border closer than 400m
d450m	indicator for slum border closer than 450m
d5	indicator for 5>= days of violence (march-november)(250m)
d500m	indicator for slum border closer than 500m
d50m	indicator for slum border closer than 50m
d5m	indicator for slum border closer than 5m
d6	indicator for 6>= days of violence (march-november)(250m)
d7	indicator for 7>= days of violence (march-november)(250m)
d750m	indicator for slum border closer than 750m
d8	indicator for 8>= days of violence (march-november)(250m)
d9	indicator for 9>= days of violence (march-november)(250m)

Table 2: Student's database dictionary (continuing)

dcontiguous	indicator for >=2 day contiguous violence Mar-Nov
dNcontiguous	indicator for >=1 day non-contiguous violence Mar-Nov
dropped	student has dropped out of school in the past (missing = -1)
dropped_class	share of students in the class that have dropped out of school in the past
dropped_m	student has dropped out of school in the past (missing = .)
girl_m	student is a girl (missing = .)
intensity_contn	number of days with conflict during academic year
kitchen	1 = school has a kitchen
lag1_d2	indicator >=2 days (t-1) Mar-Nov
lag1_d2_1trim	indicator >=2 days (t-1) 1trim
lag1_d2_2trim	indicator >=2 days (t-1) 2trim
lag1_d2_3trim	indicator >=2 days (t-1) 3trim
lag1_d2_4trim	indicator >=2 days (t-1) 4trim
lag2_d2_1trim	indicator >=2 days (t-2) 1trim
lag2_d2_2trim	indicator >=2 days (t-2) 2trim
lag2_d2_3trim	indicator >=2 days (t-2) 3trim
lag2_d2_4trim	indicator >=2 days (t-2) 4trim
lag3_d2_1trim	indicator >=2 days (t-3) 1trim
lag3_d2_2trim	indicator >=2 days (t-3) 2trim
lag3_d2_3trim	indicator >=2 days (t-3) 3trim
lag3_d2_4trim	indicator >=2 days (t-3) 4trim
language_score	student standardized language score

Table 2: Student's database dictionary (continuing)

language_score_250	student language score (Saeb scale 250/50)
lead1_d2	indicator >=2 days (t+1) Mar-Nov
lead1_d2_1trim	indicator >=2 days (t+1) 1trim
lead1_d2_2trim	indicator >=2 days (t+1) 2trim
lead1_d2_3trim	indicator >=2 days (t+1) 3trim
lead1_d2_4trim	indicator >=2 days (t+1) 4trim
math_score	student standardized math score
math_score_250	student math score (Saeb scale 250/50)
mother_educ	-1 = missing 0 = illiterate 1 = incomplete elementary 2 = elementary 3 = secondary 4 = undergraduated
mother_educ_low	student's mother is illiterate or did not finish elementary school (miss = -1)
mother_educ_low_m	student's mother is illiterate or did not finish elementary school (miss = .)
mother_higheduc	student's mother has elementary, or higher, school diploma (miss = .)
n_took_test	number of students who took the test
never_dropped	student has never dropped out of school
never_repeated	student has never failed a grade
nonwhite_class	share of nonwhite students in the class
nonwhite_m	student is not white (missing = .)
POINT_X	school latitude
POINT_Y	school longitude
principal_office	1 = school has principal's office
repeated	student has failed a grade in the past (missing = -1)

Table 2: Student's database dictionary (continuing)

repeated_class	share of students in the class that has failed a grade in the past
repeated_m	student has failed a grade in the past (missing = .)
school_lunch	1 = school offers lunch
science_lab	1 = school has a science lab
teacher_office	1 = school has teacher's office
white	student is white (missing = -1)
white_m	student is white (missing = .)
work_class	share of students in the class that work

Table 3: Teacher's database dictionary

id_school	**school identification number**
year	**year of Prova Brasil editions**
id_teacher	**teacher identification number**
age	teacher's age
changed_school	1 = teacher changed to another school in the following year
d2	indicator for >=2 days of violence (march-november)(250m)
graduated	1 = teacher has graduate degree
left_school	1 = teacher changed school in the following year or left and was not found in another school
male	1 = teacher is male
n_classes_school	number of classes in the school
n_classes_teacher	number of classes that she/he teaches (all schools)
n_teachers	number of teachers 1st-5th grade (just count profs once)
POINT_X	school latitude
POINT_Y	school longitude
undergraduated	1 = teacher has undergraduate degree
white	1 = teacher is white

Table 4: Movement's database dictionary

id_school	**school identification number**
year	**year of Prova Brasil editions**
id_student	**student identification number**
age	student's age
boy	student is a boy
d2	indicator for >=2 days of violence (march-november)(250m)
d2Ncontiguous	indicator for >=2 day non-contiguous violence Mar-Nov
d9	indicator for >=9 days of violence (march-november)(250m)
dcontiguous	indicator for >=2 day contiguous violence Mar-Nov
dropout_end	student dropped out between Dec-Mar(t+1)
dropout_within	student dropped out between Apr-Nov
fifth_grade	indicator for fifth grade
first_grade	indicator for first grade
forth_grade	indicator for forth grade
live_close	student lives within 15 min walking from school
mother_educ	-1 = missing 0 = illiterate 1 = incomplete elementary 2 = elementary 3 = secondary 4 = undergraduated
mother_educ_low	mother is illiterate or did not finished elementary school (missing = .)
nonwhite_m	student is not white (missing = .)
number_stud_school	number of students in the school
POINT_X	school latitude

Table 4: Movement's database dictionary (continuing)

POINT_Y	school longitude
second_grade	indicator for second grade
third_grade	indicator for third grade
transf_end	student transferred to another school between Dec-Mar(t+1)
transf_within	student transferred to another school between Apr-Nov
white	student is white (missing = -1)

Table 5: Favela report's database dictionary

id_favela	**favela identification number**
date	**report date (DD/MM/YYYY)**
days_gun_reports_favela	1= there is at least one report of gunfight in the favela on that date
gun_reports_favela	total number of gunfight reports per favela on that date
month	month that report was registered
total_reports_favela	total number of reports per favela on that date
year	year that report was registered

Table 6: AISP report's database dictionary

aisp	**id of the military police administrative unit**
year	**year**
days_reports	number of days with reports per year at AISP level
days_reports_year	number of days with reports per year at state level
homicides	number of homicides per year at AISP level
homicides_year	number of homicides per year at state level