Fighting Crime in Lawless Areas: Evidence from Slums in Rio de Janeiro

ONLINE APPENDIX

Christophe Bellégo and Joeffrey Drouard

February 23, 2023

Online Appendix

A Additional Descriptive Statistics

Table A.1: Descriptive statistics of UPPs

UPP	Date of BOPE intervention	Date of UPP installation	Number of police officers	Population	Number of favelas	Gang prior pacification
SantaMarta	19/11/08	19/12/08	123	4139	1	CV
Batan*	12/07/08	18/02/09	107	22176	7	contested
CidadeDeDeus	11/11/08	16/02/09	343	44515	10	CV
ChapeuMangueiraEBabilonia	11/05/09	10/06/09	107	3914	2	contested
PavaoPavaozinho	30/11/09	23/12/09	189	14062	2	CV
Tabajaras	26/12/09	14/01/10	144	8719	5	CV
Providencia	22/03/10	26/04/10	209	14765	3	CV
Borel	28/04/10	07/06/10	287	15707	6	contested
Andarai	11/06/10	28/07/10	219	14318	6	CV
Formiga	28/04/10	01/07/10	111	5036	1	CV
Salgueiro	30/07/10	17/09/10	140	4131	2	CV
Turano	10/08/10	30/10/10	173	14072	12	CV
Macacos	14/10/10	30/11/10	221	23341	2	ADA
SaoJoaoQuietoMatriz	06/01/11	31/01/11	208	9748	5	CV
CoroaFalletFogueteiro	06/01/11	25/02/11	193	14222	6	contested
EscondidinhoEPrazeres	06/01/11	25/02/11	182	9335	7	CV
Mangueira	19/06/11	03/11/11	332	17157	5	CV
SaoCarlos	06/01/11	17/05/11	244	22462	6	ADA
Vidigal	13/12/11	18/01/12	246	12452	2	ADA
Fazendinha	28/11/10	18/04/12	314	22454	5	CV
NovaBrasilia	28/11/10	18/04/12	340	33803	4	CV
AdeusBaiana	28/11/10	11/05/12	245	10606	3	CV
Alemao	28/11/10	30/05/12	320	16071	2	CV
Chatuba	27/06/12	27/06/12	230	11940	4	CV
FeSereno	27/06/12	27/06/12	170	5672	4	CV
ParqueProletario	28/11/10	28/08/12	220	17239	3	CV
VilaCruzeiro	28/11/10	28/08/12	300	19344	3	CV
Rocinha	13/12/11	20/09/12	700	71143	2	ADA
Jacarezinho	14/10/12	16/01/13	543	41903	13	CV
Manguinhos	14/10/12	16/01/13	588	24541	8	CV
AraraMandela	13/10/12	06/09/13	273	18225	5	CV
BarreiraVascoTuiuti	03/03/13	12/04/13	150	17040	4	CV
Caju	03/03/13	12/04/13	350	19411	9	ADA
CerroCora	29/04/13	03/06/13	232	3073	3	CV
CamaristaMeier	06/10/13	02/12/13	230	15290	8	CV
Lins	06/10/13	02/12/13	250	14196	10	CV
VilaKennedy	13/03/14	23/05/14	250	40606	8	CV

Notes: We consider Batan as a contested one. Prior its pacification (which starts in July 2008), Batan became controlled by militias by the end of 2007. Before, this territory was disputed between gangs and militias.

Table A.2 highlights the absence of an important correlation between the observable socioeconomic variables that characterize the UPPs. In particular, the percentage of homeowners is negatively correlated with average income per household, which might be explained by public housing and homeownership programs that were implemented in the past by the state.

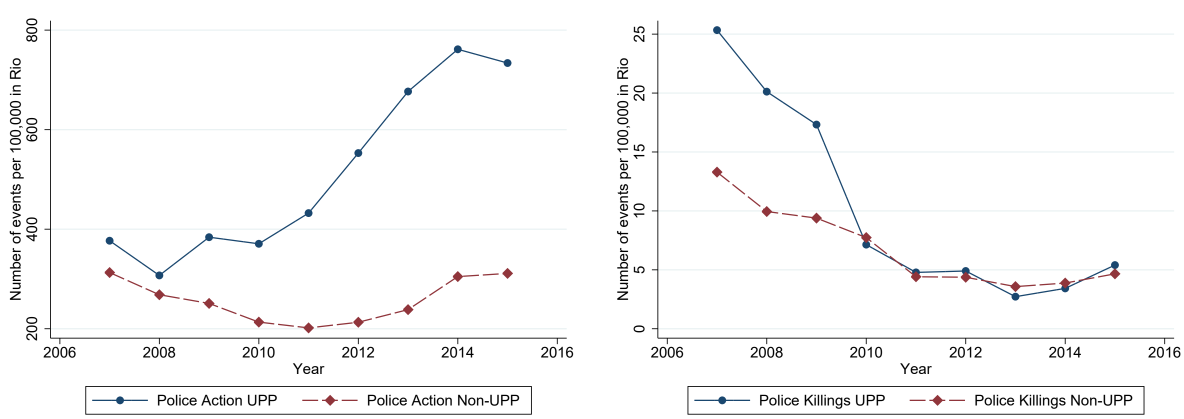
Table A.2: Correlation between socioeconomic characteristics across UPPs

	Homeowners (%)	Literacy (%)	Young (%)	Water (%)	Electricity (%)	Sewer (%)	Income (level)
Homeowners	1.0000						
Literacy	-0.3328	1.0000					
Young	0.0185	-0.1072	1.0000				
Water	-0.2443	0.0457	0.0358	1.0000			
Electricity	-0.1471	0.3428	0.1815	-0.1306	1.0000		
Sewer	-0.1411	0.2737	0.1807	0.2024	0.2294	1.0000	
Income	-0.4528	0.0340	-0.1184	0.0898	0.0326	0.2264	1.0000

Notes: Correlations calculated from the 2010 census across the 37 UPPs installed in Rio de Janeiro.

Figure A.1: Crime dynamics in UPPs and in the rest of the city (outside UPPs)

(a) Police Action (b) Police Killings



Notes: Figures plot the annual crime rates per 100,000 inhabitants in Rio de Janeiro in areas that were covered by UPPs at the end of the study period and in areas that were never covered by UPPs (*i.e.*, in the rest of the city).

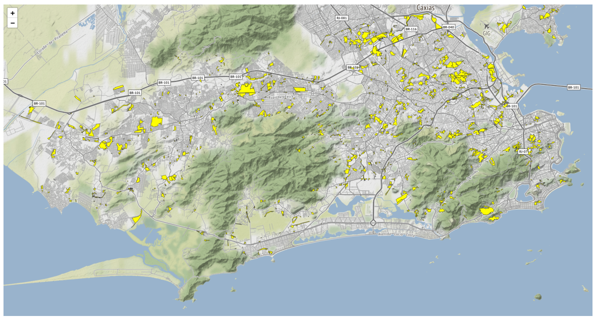
B NO CLEAR PATTERN IN THE TIMING OF PACIFICATION

As shown in Figure B.1, the geographical localization of pacified favelas over time does not follow a clear pattern. BOPE entered first two relatively big favelas, Batan and Cidade de Deus, whereas they seem to be localized in a different area of the city than the others. The public scandal following the kidnaping and torturing of a group of journalists by members of a militia led to the pacification of Batan, which is also localized close to the Olympic Games facility of Deodoro. The favelas of Cidade de Deus were then pacified because of their close proximity to the Olympic Village. Then, it seems that the State of Rio decided to pacify the small favelas close to Botafogo and Copacabana, a touristic area in the South-East of Rio de Janeiro.

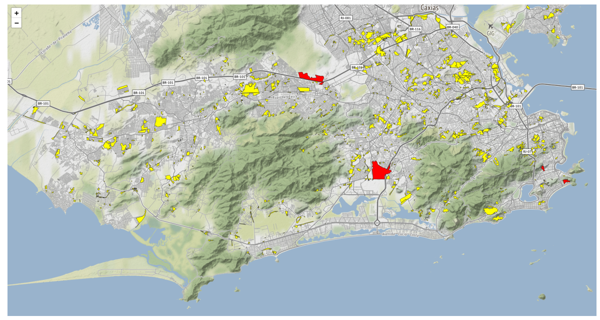
In the Figures B.2 and B.3, we plot the characteristics of the UPPs as a function of the date of pacification. Again, there is no obvious determinant of the timing of pacification.

Figure B.1: Geographic distribution of UPPs over time

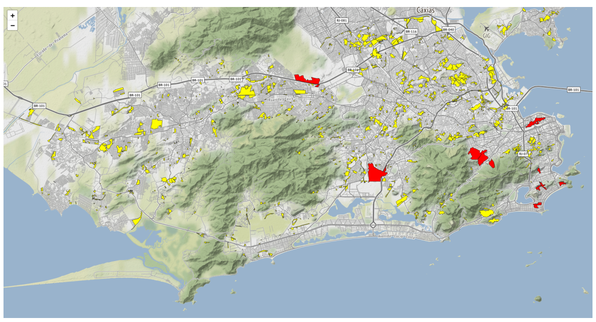
(a) June 2008



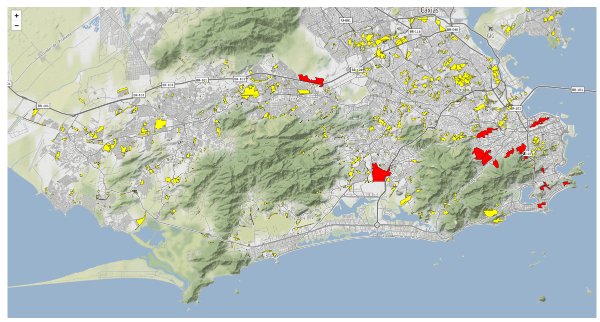
(b) March 2009



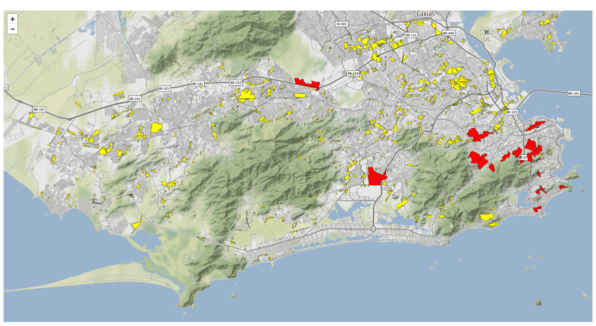
(c) August 2010



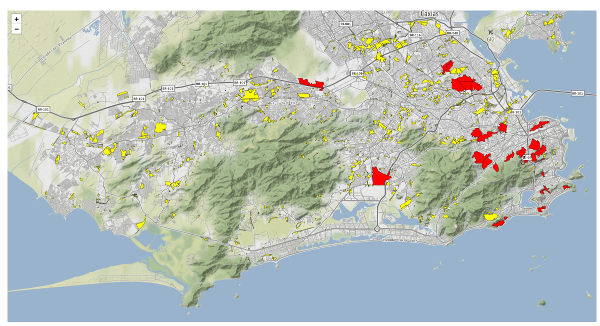
(d) December 2010



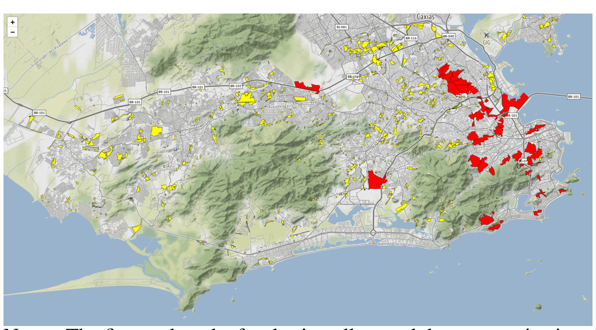
(e) June 2011



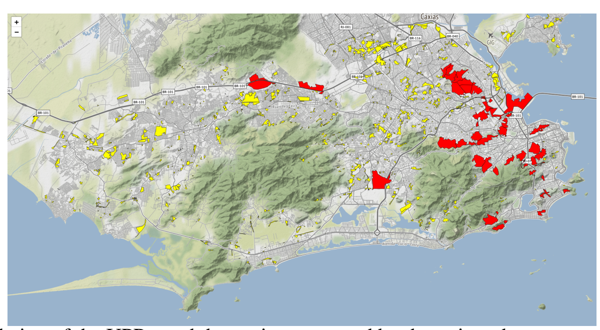
(f) July 2012



(g) September 2013



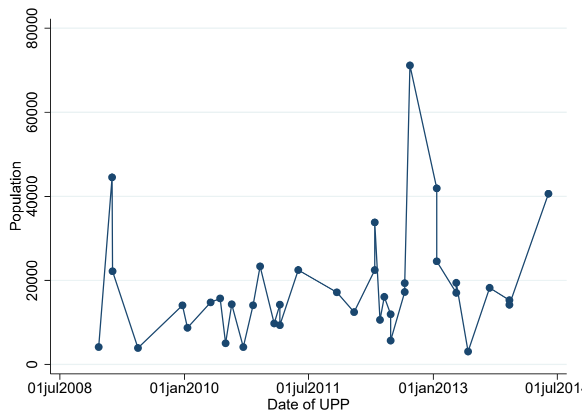
(h) June 2014



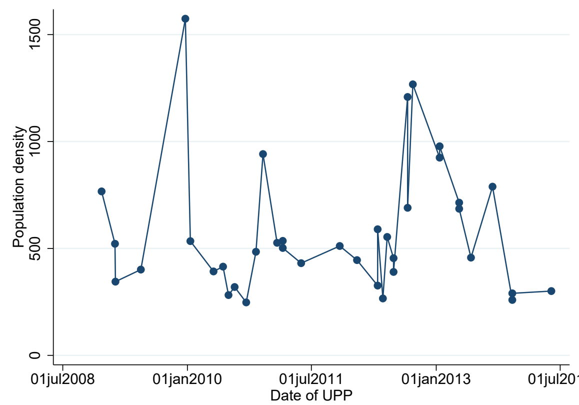
Notes: The figure plots the favelas in yellow and the progressive installation of the UPPs, and the territory covered by them, in red.

Figure B.2: Characteristics of the UPPs as a function of the date of pacification

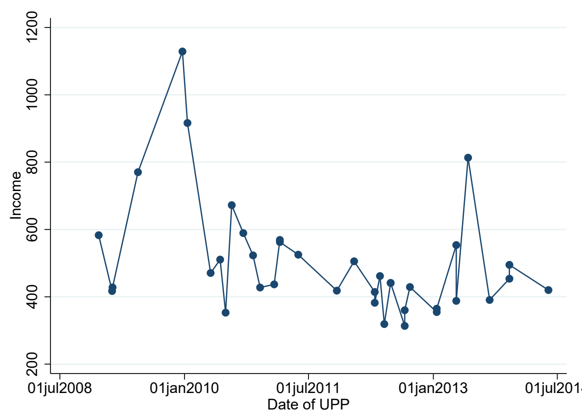
(a) Population



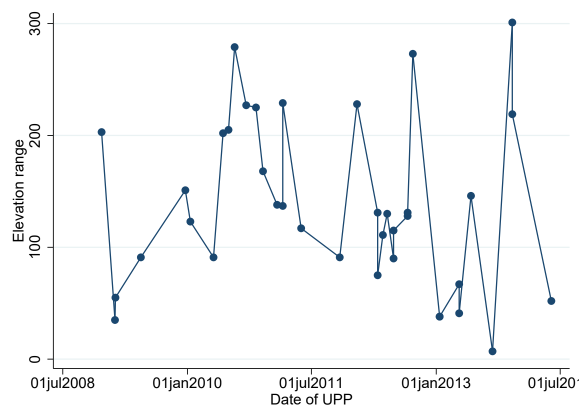
(b) Population density



(c) Income per capita



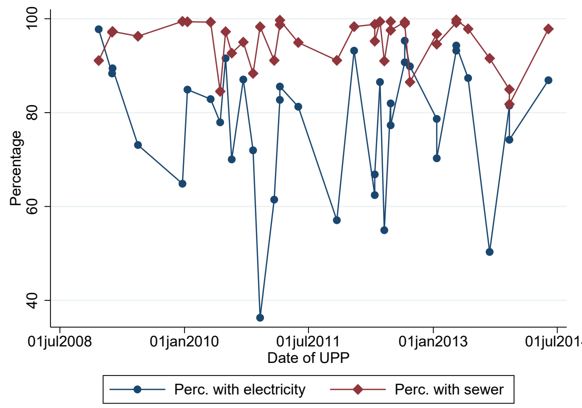
(d) Elevation range



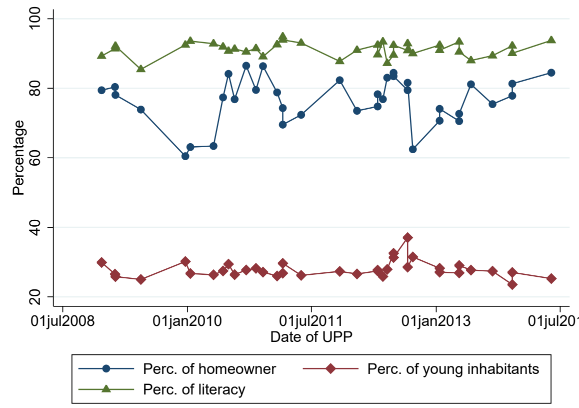
Notes: The figures plot the scatter of a given characteristic of the UPP as a function of the date of pacification.

Figure B.3: Characteristics of the UPPs as a function of the date of pacification

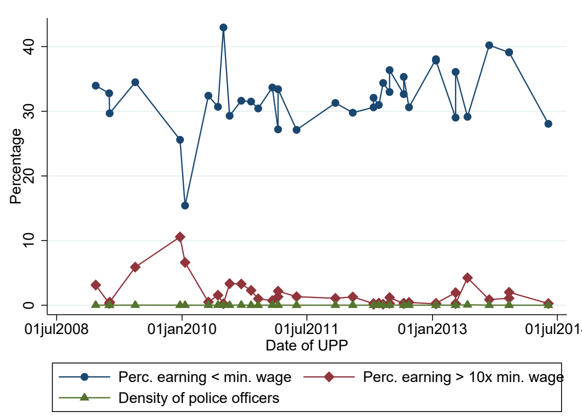
(a) Electricity and sewerage penetration



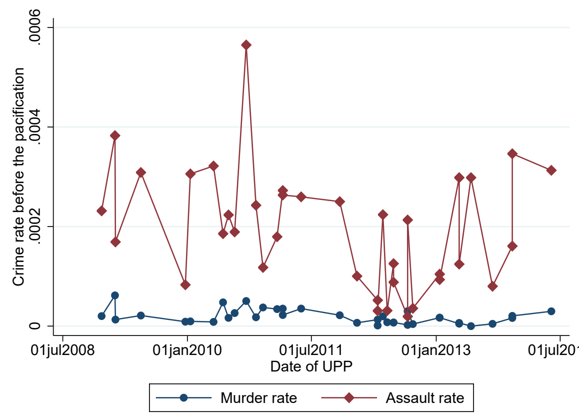
(b) Homeowner, young, and literate inhabitants



(c) Density of police officers and income inequality



(d) Murders and assaults before the pacification



Notes: The figures show the scatter of a given characteristic of the UPP as a function of the date of pacification.

C An Alternative Solution to Handle the Bias Resulting From the Unobserved Reporting Rate

The solution presented in Section III relies on two assumptions, one of them being that the policy has no direct effect on the underlying level of the proxy variable (the number of accidents). We present herein an alternative solution that relaxes this assumption by using another variable that is very close to the proxy variable but that is not affected by the variation in the reporting rate (*i.e.*, its reporting rate is 100%). Indeed, the number of accidents could increase as a result of the pacification policy if this policy led to an increase in traffic on the streets of the favelas.

The number of accidents can be decomposed in the number of fatal accidents and nonfatal accidents. We assume that the number of fatal accidents are always perfectly reported to the police, which seems quite realistic. Moreover, we allow the pacification policy to have a direct effect on the number of accidents, but we constrain this effect to be the same on fatal accidents and on nonfatal accidents. Indeed, accidents are random events, and the occurrence of fatal accidents compared to nonfatal ones is purely incidental. Thus, if the policy has an impact on accidents, it should have the same effect on fatal and nonfatal ones.

Formally,

$$\ln\left(Accident_{i,t}^{F}\right) = \chi Intervention_{i,t} + \kappa Pacified_{i,t} + X_{i,t}'\phi + d_i + d_t + u_{i,t}$$
$$\ln\left(Accident_{i,t}^{NF}\right) = \chi Intervention_{i,t} + \kappa Pacified_{i,t} + \ln(RR_{i,t}) + X_{i,t}'\lambda + c_i + c_t + e_{i,t}$$

where $Accident_{i,t}^{F}$ and $Accident_{i,t}^{NF}$ are the rate of fatal accidents and nonfatal accidents, respectively, in UPP i during month t. It is direct to obtain the following:

$$\ln\left(Accident_{i,t}^{NF}\right) - \ln\left(Accident_{i,t}^{F}\right) = \ln(RR_{i,t}) + X_{i,t}'(\lambda - \phi) + (c_i - d_i) + (c_t - d_t) + (e_{i,t} - u_{i,t}) \quad (1)$$

As before, we have the following:

$$\ln\left(Crime_{i,t}^{C,R}\right) = \alpha Intervention_{i,t} + \beta Pacified_{i,t} + \ln(RR_{i,t}) + X_{i,t}'\theta + \nu_i + \gamma_t + \varepsilon_{i,t} \quad (2)$$

By substituting $\ln(RR_{i,t})$ from equation (1) into equation (2), we obtain the following:

$$\ln\left(Crime_{i,t}^{C,R}\right) - \frac{\ln\left(Accident_{i,t}^{NF}\right)}{\ln\left(Accident_{i,t}^{F}\right)} = \alpha Intervention_{i,t} + \beta Pacified_{i,t} + X_{i,t}'\left(\theta - (\lambda - \phi)\right)$$

$$+ (\nu_i - (c_i - d_i)) + (\gamma_t - (c_t - d_t)) + (\varepsilon_{i,t} - (e_{i,t} - u_{i,t}))$$

This equation can be directly estimated by OLS to obtain an unbiased β coefficient. Table C.1 presents the results obtained with this solution. They are very similar to those obtained from the other solution and presented in Table 3. This set of findings confirms that the assumption $E[Pacified_{i,t}u_{i,t}] = 0$ in equation (5) is not a strong one.

Table C.1: Alternative solution to handle the bias resulting from the unobserved reporting rate

	Murder	Assault	Robbery	Theft	Extortion
Pacified	N.R.	0.523***	-0.329***	0.0535	-0.175**
		(0.101)	(0.0948)	(0.0790)	(0.0733)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	N.R.	N.R.	0.615***	-0.0620	0.400***
			(0.115)	(0.0681)	(0.0871)
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained for different crime indicators as the outcome variable, one in each column. The alternative correction is not applied to murders, police actions and police killing as they are assumed to be 100% reported. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. N.R. stands for Not Relevant as the correction of the reporting bias is not relevant for some crime indicators. Standard errors clustered by UPP in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

This alternative solution to handle the bias is simply a variation of what is presented in the body of the paper. This alternative solution could be useful for other papers facing a problem where the proxy variable is probably affected by the treatment but where a twin of the proxy variable exists and is unconcerned by the reporting bias.

D ADDITIONAL DETAILS ON THE BOUNDED VARIATION ASSUMPTIONS

The first solution to correct the issue from the unobserved increase in the reporting rate relies on the use of a proxy variable. It also assumes that the UPP-specific time-varying part of the reporting rate of each crime indicator is affected in the same proportion by the treatment and is equal to that of accidents. That is, $RR_{i,t}^{(C)} = RR_{i,t}^{(A)}$, $\forall C$, where $RR_{i,t}^{(j)}$ is the time-varying part of the reporting rate of event j. We here relax this assumption and assume that $RR_{i,t}^{(C)} = \rho_{i,t} RR_{i,t}^{(A)}$. When $\rho_{i,t} > 1$, there is an overreaction to the policy in the reporting rate of crime C compared to the one of accidents, and when $\rho_{i,t} < 1$, there is an underreaction.

In practice, we parametrize $\rho_{i,t}$ as follows: $\rho_{i,t} = 1 + \kappa \times Pacified_{i,t}$. The ratio of the variation rate of $RR_{i,t}^{(C)}$ to that of $RR_{i,t}^{(A)}$ writes $\Delta = ((1+\kappa)(1+\delta) - 1)/\delta$, where δ captures the increase in the reporting rate of accidents.[1] To estimate δ, we add more structure in the relation between the accident reporting rate and the treatment by specifying the following equation:

$$\ln\left(RR_{i,t}^{A}\right) = \zeta Intervention_{i,t} + \delta Pacified_{i,t} + X_{i,t}'\omega + a_i + a_t + e_{i,t} \tag{3}$$

In this equation, the δ parameter identifies the effect of pacification on the accident reporting rate. We know that

$$\ln\left(Accident_{i,t}^{R}\right) = X_{i,t}'\lambda + d_i + d_t + \ln\left(RR_{i,t}^{A}\right) + u_{i,t} \tag{4}$$

By substituting the expression of $\ln(RR_{i,t}^{A})$ that we obtain from equation (3) into equation (4), we obtain the following:

$$\ln\left(Accident_{i,t}^{R}\right) = \zeta Intervention_{i,t} + \delta Pacified_{i,t} + X_{i,t}'(\lambda + \omega) + (d_i + a_i) + (d_t + a_t) + (u_{i,t} + e_{i,t}) \tag{5}$$

Consequently, the increase in the accident reporting rate δ is directly identified by the estimation of equation (5).

[1] Consider two periods, one before (t_0) and one after (t_1) the pacification. The time-varying part of the reporting rate is the same for all events before the pacification: $RR_{i,t_0}^{(C)} = RR_{i,t_0}^{(A)}$. After the pacification, we have $RR_{i,t_1}^{(C)} = (1+\kappa) RR_{i,t_1}^{(A)}$. Knowing that $RR_{i,t_1}^{(A)} = (1+\delta) RR_{i,t_0}^{(A)}$, it is straightforward to show that $\left[\left(RR_{i,t_1}^{(C)} - RR_{i,t_0}^{(C)}\right)/RR_{i,t_0}^{(C)}\right] / \left[\left(RR_{i,t_1}^{(A)} - RR_{i,t_0}^{(A)}\right)/RR_{i,t_0}^{(A)}\right] = \left[(1+\kappa)(1+\delta) - 1\right]/\delta$.

By varying the parameter κ, we simulate different values of Δ, the variation of $RR_{i,t}^{(C)}$ relative to that of $RR_{i,t}^{(A)}$, and we obtain bounds of the effects. In particular, we choose values of κ such that the ratio Δ takes value every 0.25 in the interval $[0;5]$. The interval is large so that we can identify the relative increase in the reporting rate that would lead to a reversal of the effect. For instance, when $\Delta = 1$, we have $RR_{i,t}^{(C)} = RR_{i,t}^{(A)}$, $RR_{i,t}^{(C)}$ increases as much as $RR_{i,t}^{(A)}$ so that we point-identify the coefficient β as before. When $\Delta = 2$, $RR_{i,t}^{(C)}$ increases twice as much as $RR_{i,t}^{(A)}$. When $\Delta = 0$, $RR_{i,t}^{(C)}$ is left unchanged by the policy even though $RR_{i,t}^{(A)}$ increases.

E Optimal Value of the Constant Added to All Crime Observations in the Log-Regressions

The occurrence of events such as murders is, fortunately, relatively rare. Crime data available at a detailed geographical level generally contain many zeros, such that the use of a logarithm function might be problematic. Despite this difficulty, we employ a log-linear specification for several reasons. First, the empirical model naturally writes in log. It allows us to deal with the reported nature of crime data naturally and to estimate the increase in the reporting rate that is implied by the policy. Second, the use of a log specification to study crime is standard in the literature, so our estimation results are directly comparable to those of other studies (see, for instance, Levitt, 1998; Ayres and Levitt, 1998; Draca et al., 2011). Third, we believe that the effects are more likely to be multiplicative than additive. Last and not least, we are able to identify the size of the increase in the reporting rate with a log specification. It is harder to estimate the omitted variable bias with a Poisson regression for instance. Therefore, a specification in log seems to be more appropriate than other standard ones (*i.e.*, OLS in level or Poisson regression), provided that we carefully handle the problem of zeros.

To address this issue, we add a small constant to all crime data points, as $\log(0)$ is undefined. The choice of the constant value is key to minimize the bias that it will mechanically introduce. In general, adding the smallest possible value is not the best solution, as it can change the distribution of the data, depending on the value of the observations (Bellégo et al., 2022). Noting that a log transformation squeezes high values and expands low values, the objective is to add a constant that tries to preserve the initial order of magnitude in the data and that approximately maps zero to zero. A rule of thumb is to add a constant that is close to the lowest strictly positive observation. For instance, with crime data, the lowest value above zero is one; thus, it is advised to add 0.5 or 1 to all the data points before applying the log function.

Formally, we begin by testing the sensitivity of the results to the choice of the constant that is added to all data points by varying this parameter near 0.5, with $c = \{0.25, 0.5, 1\}$. The results are presented in Table E.1, in panels A, B, and C. These results show that the sign of the effect of the pacification is not affected by the choice of the constant, while the magnitude of the effect differs across the different constants.

Then, we look for the constant parameter that provides the closest results to what we obtain when estimating the model with OLS in level (*i.e.*, without the log transformation). We first divide the coefficient estimated from the OLS in level by the mean value of the crime rate over the period. We then compare this effect to the coefficients obtained from an OLS regression in log when adding a constant c, with $c = \{0.005, 0.01, 0.05, 0.1, 0.25, 0.5, 1\}$. We conduct this test for the assault indicator without correcting for the unobserved reporting rate. The results are presented in Table E.2. The effect obtained from the OLS specification without log (column 1) represents approximately 0.70% of the mean value of the assault rate, which is closest to the effect obtained with a log regression when adding a constant c equal to 0.5 (column 3). Therefore, in the paper, we will always add a constant $c = 0.5$ to all crime observations.

Table E.1: Estimates testing different constants without bias correction

Panel A. Adding a constant $c = 1$ to all data points					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.0487*	0.621***	-0.0984*	0.196***	0.0102
	(0.0257)	(0.0828)	(0.0534)	(0.0558)	(0.0109)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	0.631***	-0.112***	0.654***	0.0827***	0.569***
	(0.110)	(0.0298)	(0.0788)	(0.0233)	(0.0620)
Panel B. Adding a constant $c = 0.5$ to all data points					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.0688*	0.715***	-0.138**	0.245***	0.0157
	(0.0340)	(0.0943)	(0.0653)	(0.0671)	(0.0170)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	0.751***	-0.165***	0.806***	0.129***	0.591***
	(0.129)	(0.0424)	(0.0925)	(0.0340)	(0.0662)
Panel C. Adding a constant $c = 0.25$ to all data points					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.0919**	0.793***	-0.182**	0.291***	0.0225
	(0.0433)	(0.105)	(0.0786)	(0.0794)	(0.0246)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	0.866***	-0.228***	0.955***	0.187***	0.605***
	(0.148)	(0.0568)	(0.107)	(0.0467)	(0.0695)
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trend	Yes	Yes	Yes	Yes	No
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (1) for different crime indicators as the outcome variable, one in each column. We add a constant $c = 1$, $c = 0.5$ and $c = 0.25$ to all crime observations in Panel A, B and C, respectively. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP$\times$month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.2: Comparing different constants used in log-regressions with OLS in level for assaults

	Outcome = Assault rate (mean value = 0.0003923)							
	Regression in level	Regressions in log						
	no constant	$c = 1$	$c = 0.5$	$c = 0.25$	$c = 0.1$	$c = 0.05$	$c = 0.01$	$c = 0.005$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pacified	0.000299***	0.621***	0.715***	0.793***	0.881***	0.940***	1.067***	1.120***
	(0.0000440)	(0.0828)	(0.0943)	(0.105)	(0.120)	(0.131)	(0.160)	(0.173)
Intervention	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218	4218	4218	4218

The table presents the treatment effect obtained from the estimation of equation (1) with OLS in level (*i.e.*, without the log transformation) in column (1). Then, we present the coefficients estimated from equation (1) when adding a constant c to all observations, with $c = \{0.005, 0.01, 0.05, 0.1, 0.25, 0.5, 1\}$ in columns (2) to (8), respectively. We only use the assault rate as the outcome variable. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F Main results without UPP linear time trends

Table F.1: Baseline results with or without the correction

Panel A. Without correction of the reporting bias					
	Murder	Assault	Robbery	Theft	Extortion
Intervention	-0.00461	0.557***	0.0417	0.325***	0.0341
	(0.0470)	(0.126)	(0.121)	(0.0976)	(0.0259)
Pacified	-0.0516*	0.738***	-0.106	0.249***	0.00993
	(0.0255)	(0.0937)	(0.0779)	(0.0735)	(0.0168)
	Police Action	Police Killings	Threats	Rape	Total Events
Intervention	0.841***	-0.0753	0.708***	0.125**	0.598***
	(0.161)	(0.0460)	(0.128)	(0.0517)	(0.101)
Pacified	0.780***	-0.146***	0.793***	0.118***	0.632***
	(0.128)	(0.0337)	(0.0927)	(0.0350)	(0.0705)
Panel B. With correction of the reporting bias					
	Murder	Assault	Robbery	Theft	Extortion
Intervention	N.R.	0.486***	-0.0294	0.254**	-0.0370
		(0.129)	(0.154)	(0.114)	(0.0661)
Pacified	N.R.	0.494***	-0.350***	0.00555	-0.234***
		(0.0986)	(0.0940)	(0.0741)	(0.0557)
	Police Action	Police Killings	Threats	Rape	Total Events
Intervention	N.R.	N.R.	0.637***	0.0542	0.527***
			(0.123)	(0.0834)	(0.116)
Pacified	N.R.	N.R.	0.549***	-0.126**	0.388***
			(0.108)	(0.0616)	(0.0827)
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	No	No	No	No	No
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (1) in Panel A (no correction) and of equation (6) in Panel B (with the correction based on the proxy variable, which is found to increase by 23% following pacification) for different crime indicators as the outcome variable, one in each column. The correction is not applied to murders, police actions and police killings as they are assumed to be 100% reported. All the regressions include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. N.R. stands for Not Relevant as the correction of the reporting bias is not relevant for some crime indicators. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G HETEROGENEITY OF THE EFFECT

We study several socioeconomic characteristics that are likely to influence the causal effect of the pacification policy (the percentage of homeowners, of literacy, and of young people aged between 14 and 30, the population density, the altitude range, and the number of deployed police officers).[2] The variables obtained from the census are observed in 2010, at the beginning of the implementation of the policy (just three favelas were pacified before 2010), which prevents the pacification policy from generating important endogenous variations of these variables. In this analysis, we are not interested in the direct causal effect of these characteristics on crime, which would be quite challenging to identify. Instead, we analyze their interaction effect with the treatment, controlling for unobserved fixed heterogeneity that could generate low or high values for these variables.

Some interaction terms between these variables and the treatment may be endogenous if they were considered separately (*e.g.*, investigating the interaction effect with homeownership without controlling for the interactions with other variables, such as income, could generate an omitted variable bias). For this reason, it is important to simultaneously estimate several interaction effects. In summary, we do not claim to present the causal effects of these heterogeneous effects but instead document interesting correlations that are well controlled for.

Table G.1 presents the interaction effects estimated for the main crime indicators. Homeownership, literacy, and average income apparently improve the efficiency of the policy. The negative effect of education on crime occurrence has been demonstrated in Lochner and Moretti (2004). We provide new evidence that supports this mechanism at the most basic level of education, *i.e.*, the murder rate decreases more in favelas where the inhabitants are more literate. Individuals that are more educated may understand better that it is in their interest to react positively to the pacification policy. We also find that the murder rate declines more in favelas with more homeowners. It has been shown that homeowners develop stronger links with their neighbors (DiPasquale and Glaeser, 1999), but very few studies demonstrate convincing evidence about the effect of homeownership on crime (anecdotical evidence is found in Glaeser and Sacerdote, 1999). Homeowners may watch their neighborhood more closely following pacification, which could prevent some crimes from taking place. Last, the pacification policy reduced murders more in favelas where the income per capita is higher. The relation between crime and income is ambiguous, as shown in Ehrlich (1973), because higher income can

[2]We collect altitude data from the Shuttle Radar Topography Mission (see Jarvis et al., 2008).

Table G.1: Heterogeneous effects

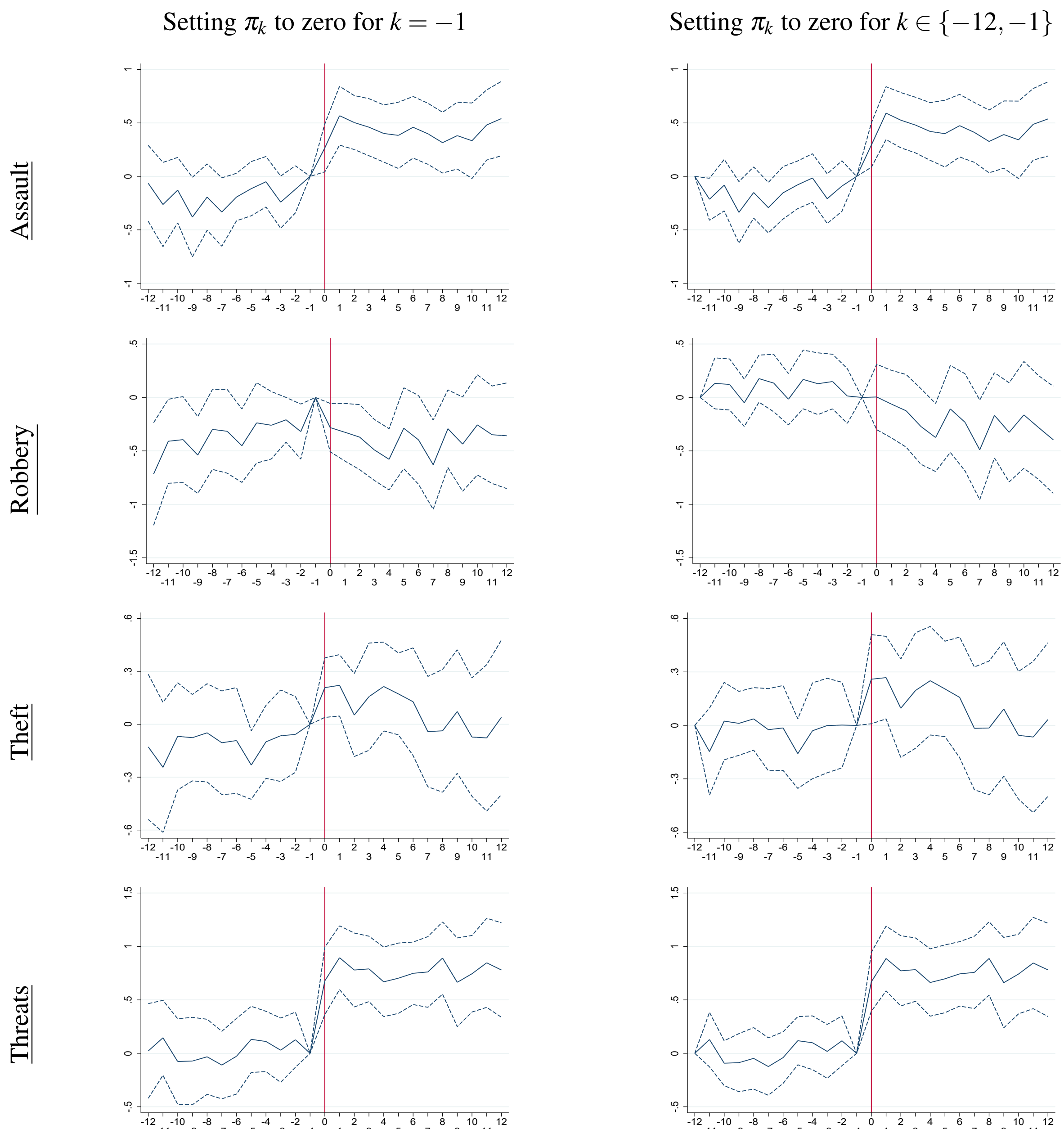
	Murder	Assault	Rape	Robbery	Theft
Pacified	2.831**	-1.331	2.624	9.822**	3.302
	(1.326)	(4.761)	(2.650)	(3.716)	(3.601)
Pacified × Homeowner	-0.0172***	0.0197	0.000140	-0.0311**	-0.0109
	(0.00578)	(0.0243)	(0.0119)	(0.0149)	(0.0175)
Pacified × Literacy	-0.0187*	0.00197	-0.0300	-0.0833***	-0.0307
	(0.0100)	(0.0334)	(0.0228)	(0.0306)	(0.0308)
Pacified × Income	-0.000301*	0.000338	0.000184	-0.000401	0.000198
	(0.000159)	(0.000892)	(0.000471)	(0.000568)	(0.000848)
Pacified × Youth	0.0127	-0.0247	0.00142	0.0255	0.0129
	(0.0105)	(0.0443)	(0.0336)	(0.0301)	(0.0488)
Pacified × Altitude range	0.000599*	0.000837	0.000146	-0.000555	-0.0000113
	(0.000354)	(0.000890)	(0.000655)	(0.00100)	(0.000967)
Pacified × Police officers	-0.000440	0.000958	0.0000312	-0.000777	0.000373
	(0.000324)	(0.00114)	(0.000743)	(0.000721)	(0.000906)
Pacified × Pop. density	-0.0000899	0.000550	-0.000247	-0.000759***	-0.000327
	(0.000114)	(0.000440)	(0.000336)	(0.000276)	(0.000453)
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes [2]
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the interaction effects, between socioeconomic characteristics and the treatment, estimated for five different crime indicators as the outcome variable, one in each column. The UPPs' socioeconomic characteristics are obtained from the 2010 census. We do not correct for the reporting bias for Murder. All the regressions take into account the intervention date and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

increase the opportunity cost of committing a crime, but it can also increase the wealth to be stolen, with the direction of the global effect depending mainly on the degree of risk aversion. Our empirical findings are more in line with the former explanation, confirming the idea that poverty is intrinsically linked to crime.

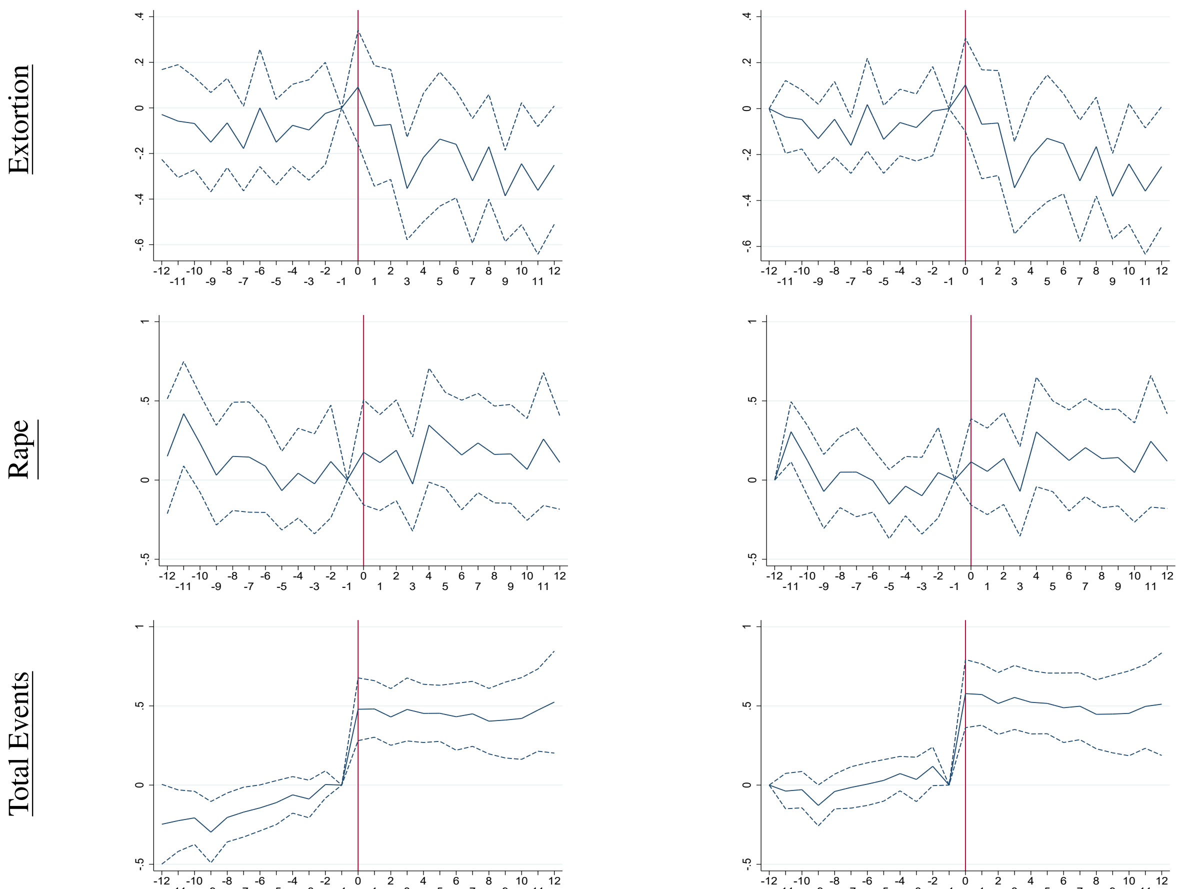
H Robustness Checks on the Event Studies

Figure H.1: Event studies for different crime indicators with a 23% increase in reporting (1/2)



Notes: The figures plot the quarterly crime rate, for different crime indicators, as a function of time since BOPE intervention. The reporting of each crime indicator is assumed to increase by 23% once the territory is pacified. The solid lines correspond to the values of π_k ($k < 0$) and τ_k ($k \geq 0$), as a function of k, obtained from the estimation of equation (7) on the sample of all UPPs for which $k \in [-12, +12]$. Standard errors are clustered at the UPP level and dashed lines represent the 95% confidence interval.

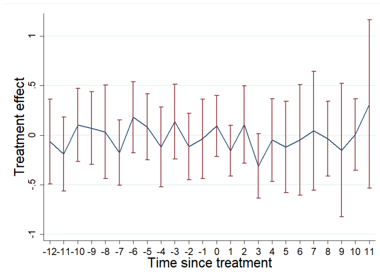
Figure H.2: Event studies for different crime indicators with a 23% increase in reporting (2/2)



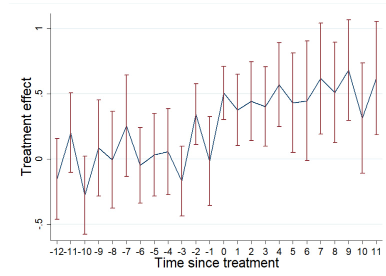
Notes: The figures plot the quarterly crime rate, for different crime indicators, as a function of time since BOPE intervention. The reporting of each crime indicator is assumed to increase by 20% once the territory is pacified. The solid lines correspond to the values of π_k ($k < 0$) and τ_k ($k \geq 0$), as a function of k, obtained from the estimation of equation (7) on the sample of all UPPs for which $k \in [-12, +12]$. Standard errors are clustered at the UPP level and dashed lines represent the 95% confidence interval.

Figure H.3: Event studies with heterogenous treatment effects following de Chaisemartin and D'Haultfœuille (2020)

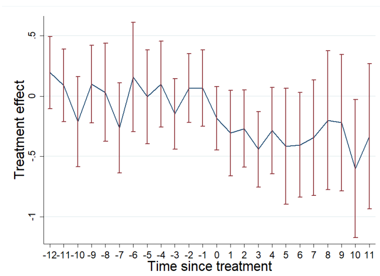
(a) Murder



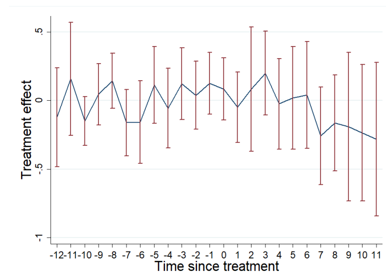
(b) Assault



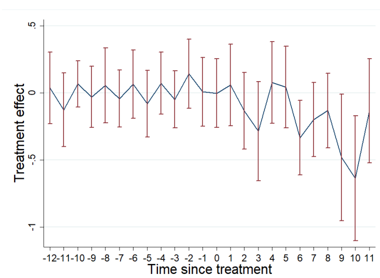
(c) Robbery



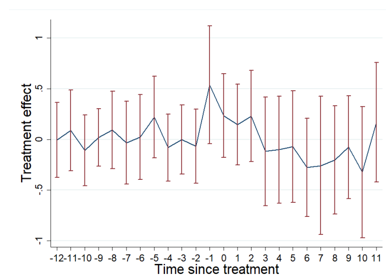
(d) Theft



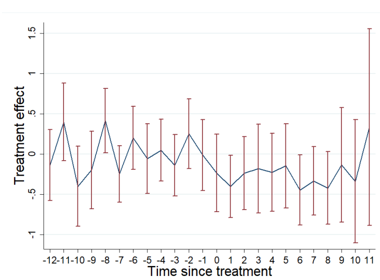
(e) Extortion



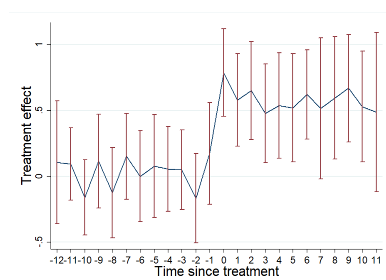
(f) Police Action



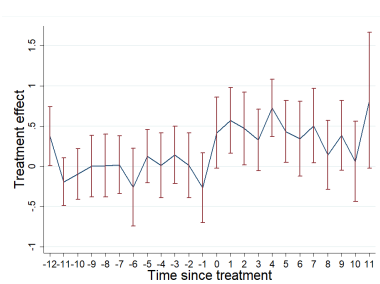
(g) Police Killings



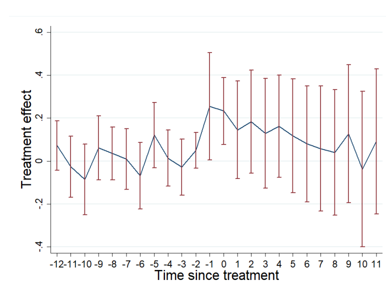
(h) Threats



(i) Rape



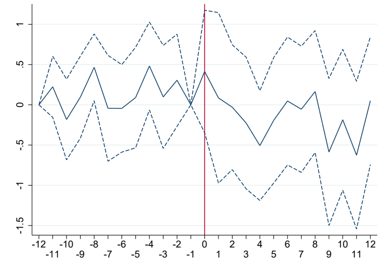
(j) Total Events



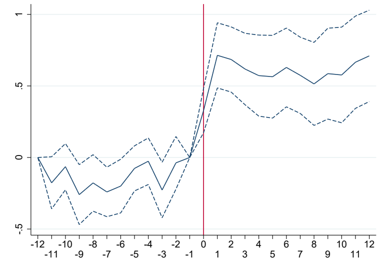
Notes: The figures plot the quarterly crime rate, for different crime indicators, as a function of the time since BOPE intervention. The solid lines correspond to the estimated dynamic treatment effect using the method of de Chaisemartin and D'Haultfœuille (2020). Standard errors are clustered at the UPP level and red intervals represent the 95% confidence interval.

Figure H.4: Event studies with Negative Binomial regressions

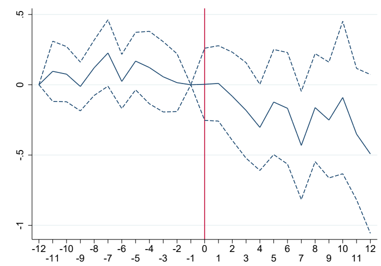
(a) Murder



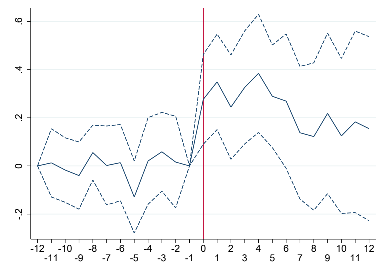
(b) Assault



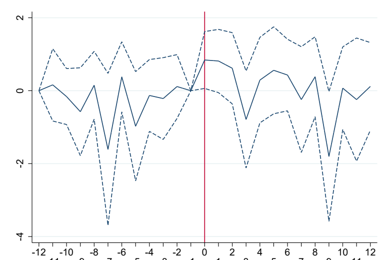
(c) Robbery



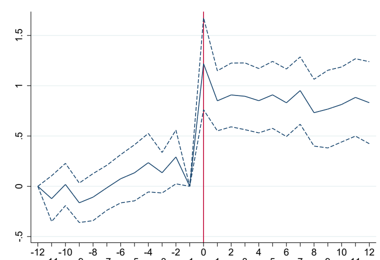
(d) Theft



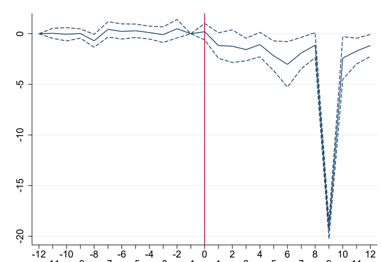
(e) Extortion



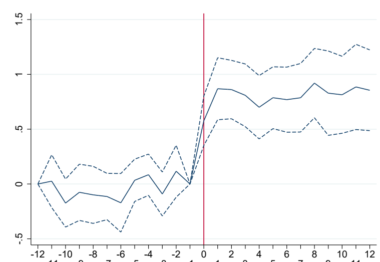
(f) Police Action



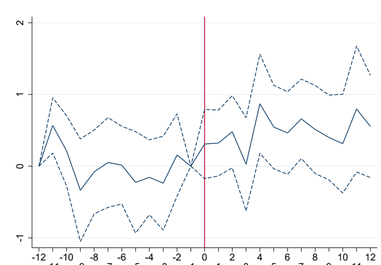
(g) Police Killings



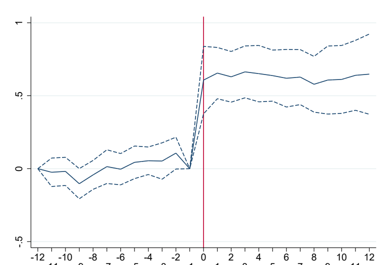
(h) Threats



(i) Rape



(j) Total Events



Notes: The figures plot the quarterly crime rate, for different crime indicators, as a function of the time since BOPE intervention. The solid lines correspond to the estimated dynamic treatment effect using a Negative Binomial regression. The specifications additionally restrict the event-quarter dummy $k = -12$ to be zero. Standard errors are clustered at the UPP level and dashed lines represent the 95% confidence interval.

I TESTS OF SPILLOVER EFFECTS BETWEEN FAVELAS

The pacification of the headquarter of CV, located in Complexo do Alemão and Complexo da Penha, started in November 2010.[3] We compare the crime rates in favelas whose pacification has not yet begun, controlled either by CV or rival criminal factions, before and after the date the BOPE entered the headquarter of CV. To implement this test, we estimate the following equation:

$$\ln\left(Crime_{i,t}^{C,R}\right) = \nu_i + \gamma_t + \rho CV_i \times CVHQ\_Pacification_t + X_{i,t}'\theta + \varepsilon_{i,t} \tag{6}$$

where CV_i indicates whether the UPP i was controlled by CV before pacification, and $CVHQ\_Pacification_t$ indicates that the pacification of the headquarter of CV has started (*i.e.*, the BOPE has entered it). We focus on the potential short-term effects that could originate from this shock, one year after it occurred, to avoid mixing them with the effects of the forthcoming pacifications, so we drop observations after 2011. We keep only the favelas that BOPE had not started to pacify at the end of 2011, which leaves us with 11 UPPs that are used in this analysis.[4] These favelas were all yet to be pacified; therefore, there is no reason for the reporting rate to vary with the treatment variable. Table I.1 presents the estimated coefficient from equation (6). No clear spillover effect between favelas is associated with the pacification of the headquarter of CV.

Table I.1: Spillover effects between favelas following the pacification of CV's headquarter

	Murder	Assault	Robbery	Theft	Extortion
CV × CVHQ_Pacification	0.0614	0.169	-0.0736	0.123	-0.00195
	(0.157)	(0.276)	(0.178)	(0.137)	(0.0575)
	Police Action	Police Killings	Threats	Rape	Total Events
CV × CVHQ_Pacification	0.0819	0.118	0.162	0.00999	-0.0746
	(0.422)	(0.0860)	(0.253)	(0.133)	(0.114)
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	660	660	660	660	660

Notes: The table presents the estimation of the parameter ρ in equation (6) for different crime indicators as the outcome variable, one in each column. All the regressions include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. We drop observations after 2011. Standard errors clustered by UPP in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[3] Complexo do Alemão and Complexo da Penha contain the following UPPs: Fazendinha, Nova Brasilia, Adeus Baiana, Alemao, Chatuba, Fe Sereno, Parque Proletario and Vila Cruzeiro. Apart from Chatuba and Fe Sereno, BOPE entered these territories in November 2010.

[4] BOPE entered Chatuba and Fe Sereno in June 2012. As they are inside Complexo da Penha, we did not keep Chatuba and Fe Sereno. Results are similar if we include these two UPPs.

Table I.2: Two more tests showing the absence of spillover effects between favelas

Panel A. Without the UPPs containing the gangs headquarters					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.0923**	0.506***	-0.379***	0.0497	-0.207***
	(0.0375)	(0.103)	(0.0935)	(0.0828)	(0.0662)
Bias Correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	0.767***	-0.197***	0.541***	-0.0805	0.371***
	(0.132)	(0.0477)	(0.118)	(0.0636)	(0.0847)
Bias Correction	No	No	Yes	Yes	Yes
Panel B. Without the UPPs containing contested favelas					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.0645	0.463***	-0.347***	0.0156	-0.240***
	(0.0416)	(0.113)	(0.0881)	(0.0831)	(0.0562)
Bias correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threats	Rape	Total Events
Pacified	0.751***	-0.177***	0.516***	-0.111*	0.338***
	(0.145)	(0.0453)	(0.117)	(0.0623)	(0.0921)
Bias correction	No	No	Yes	Yes	Yes
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations (Panel A)	3192	3192	3192	3192	3192
Observations (Panel B)	3762	3762	3762	3762	3762

Notes: The table presents the treatment effects obtained from the estimation of equation (6) for different crime indicators, excluding some UPPs. When a UPP is excluded, it is for the entire period between January 2007 and June 2016. We do not correct for the reporting bias for Murder, Police Action and Police Killings. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP$\times$month. In Panel A, the sample excludes UPPs that contain gang headquarters, i.e., Fazendinha, Nova Brasilia, Adeus Baiana, Alemao, Chatuba, Fe Sereno, Parque Proletario and Vila Cruzeiro for CV, and Rocinha for ADA. In Panel B, the sample excludes UPPs that were previously contested, i.e., Batan, Chapeu Mangueira Babilonia, Borel, Coroa Fallet Fogueteiro. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

J Alternative Treatment Variable

We show robustness to another construction of the pacification variable. In this test, we define the treatment associated to pacification as the period starting from the date of BOPE intervention, instead of the date of pacification. Table J.1 presents the findings obtained with this alternative treatment variable and they are similar to those obtained in Table 3.

Table J.1: Alternative Construction of the Pacification Variable

	Murder	Assault	Robbery	Theft	Extortion
BOPE_UPP	-0.0573	0.516***	-0.264***	0.104	-0.137**
	(0.0346)	(0.0967)	(0.0839)	(0.0696)	(0.0543)
Bias correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threats	Rape	Total Events
BOPE_UPP	0.766***	-0.147***	0.627***	-0.0269	0.424***
	(0.121)	(0.0347)	(0.105)	(0.0514)	(0.0769)
Bias correction	No	No	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (6) for different crime indicators as the outcome variable, with an alternative treatment variable, BOPE_UPP, that is equal to the sum of *Intervention* variable and *Pacified* variable. We do not correct for the reporting bias for Murder, Police Action and Police Killings. All the regressions include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

K Robust Inference

To carry out a correct inference with clustered standard errors, it is necessary to have a sufficient number of clusters. In practice, having 30 to 40 clusters is usually considered the minimum number of clusters for the asymptotic property of the Wald statistic to be valid. In our application, there are 37 clusters (UPPs), which is just at the limit of being sufficient, so that we might underestimate the standard errors. Therefore, we implement two different procedures to correct this issue. First, we implement the wild cluster bootstrap procedure proposed in Cameron et al. (2008), which is supposed to perform well with a very limited number of clusters. The p-values obtained with this procedure are presented in Table K.1, and are in line with the main results displayed in Table 3.

Second, we run a randomization test, following Fisher (1935), which does not rely on asymptotic properties. In a standard randomization test, the attribution of the treatment is randomized between

Table K.1: A wild cluster bootstrap procedure for the inference of main effects

	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.0688**	0.509***	-0.344***	0.0387	-0.190
P-value	(0.0350)	(0.0010)	(0.0000)	(0.6210)	(0.0040)
Bias correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	0.751***	-0.165***	0.601***	-0.0767	0.385***
P-value	(0.0000)	(0.0000)	(0.0000)	(0.2080)	(0.0000)
Bias correction	No	No	Yes	Yes	Yes
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (6) for different crime indicators as the outcome variable, one in each column. We do not correct for the reporting bias for Murder, Police Action and Police Killings. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP$\times$month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. We implement the wild cluster bootstrap procedure to estimate standard errors. P-values in parentheses. They are obtained from 1000 replications of the wild cluster bootstrap procedure. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the treated and the untreated groups. In our case, all the groups (UPPs) received the treatment, so we cannot randomize on this dimension. To adapt the test, we randomized the treatment date of each UPP. The dates of the BOPE's intervention are between July 2008 and March 2014, and the pacification dates are between December 2008 and May 2014. Since the minimum duration of the intervention is equal less than one month, we attribute to each UPP a (uniform) random pacification date between July 2008 and May 2014. We do not randomize the intervention durations. A randomized intervention date is calculated so that it is equal to the randomized treatment date minus the intervention duration. Thus, we implicitly assume that the intervention duration that is specific to each UPP, stems from the UPP's characteristics and does not depend on the date of pacification. To test this assumption, we model the intervention duration as a linear function of some UPPs' characteristics and the rank of pacification. Table K.2 shows that the intervention duration is determined by the UPP's characteristics but not by the timing of pacification, which supports our procedure.

With this randomization test, we test the null hypothesis that the average treatment effect is zero. If the null hypothesis is true, the increase (or decrease) in crime in each UPP will be the same regardless of when the treatment is received. Therefore, the observed test statistic (*i.e.*, the real estimated treatment effect) should not differ much from all the randomized test statistics. It is then easy to compute the p-value as the proportion of test statistics that are superior (in absolute value) to the observed

Table K.2: Determinants of the UPP's intervention duration

	Intervention duration		
Rank of pacification	-0.00482		-0.0707
	(0.114)		(0.136)
Average income per capita		-0.0132***	-0.0138***
		(0.00448)	(0.00498)
Number of favelas		-1.055**	-1.024**
		(0.389)	(0.400)
Population		0.000551**	0.000541**
		(0.000241)	(0.000248)
Population²		-6.84e-09**	-6.70e-09**
		(3.16e-09)	(3.27e-09)
Constant	5.167***	10.51**	11.66**
	(1.770)	(4.170)	(5.149)
Observations	37	37	37
R-squared	0.0000151	0.369	0.371

Notes: Estimation of the intervention duration as a linear function of some UPPs' characteristics and the rank of pacification. Average income per capita and Population are obtained from the 2010 census. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

test statistic. Table K.3 presents the p-values obtained from 1000 permutations of the treatment date and confirms the robustness of the results.

Table K.3: Randomization test of the pacification date for inference of the main effects

	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.069*	0.509***	-0.344***	0.039	-0.190***
P-value	(0.057)	(0.001)	(0.001)	(0.606)	(0.003)
Bias correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	0.751***	-0.165***	0.601***	-0.077	0.385***
P-value	(0.001)	(0.001)	(0.001)	(0.238)	(0.001)
Bias correction	No	No	Yes	Yes	Yes
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (6) for different crime indicators as the outcome variable, one in each column. We do not correct for the reporting bias for Murder, Police Action and Police Killings. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. We implement the randomization inference test described above to estimate standard errors by randomizing the treatment date of each UPP. P-values in parentheses. They are obtained from 1000 permutations of the treatment date. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

L TAMPERING WITH DATA

A possible concern regarding the use of official crime data is that these data can be manipulated by the police. When we do not correct for reporting bias, our estimations highlight negative effects on outcomes such as murders or robberies, but it also produces important positive effects on assaults, rapes, thefts and threats. Overall, we find an important increase in the total number of events associated with the policy. If these official data were manipulated by the police, it would be unlikely that we would observe such a strong increase in any crime category.

Furthermore, we decompose the pacification effect among the twelve months of the year. This decomposition allows us to check whether the number of reported crimes reacts more to pacification at the end of the year (November-December) or at the end of each trimester, when the police could falsify the numbers to virtually reach a crime level goal, as objectives are usually fixed at a quarterly or a yearly level (see Posner, 2010, for a similar argument). The results are presented in Table L.1. They do not exhibit a clear decrease in any high-level crime or a clear increase in the action of the police at the end of the year or at the end of a quarter. Overall, the evidence is not consistent with police tampering with crime data.

Table L.1: Decomposition of the pacification effect over twelve months

	Murder	Assault	Robbery	Theft	Police Action	Police Killings	Threats	Rape
Pacif × Jan	-0.217***	0.735***	-0.345*	0.0245	0.940***	-0.142**	0.575**	0.0481
	(0.0650)	(0.188)	(0.173)	(0.144)	(0.215)	(0.0590)	(0.226)	(0.168)
Pacif × Feb	-0.183***	0.541***	-0.429***	0.0616	0.981***	-0.0838	0.559***	-0.161
	(0.0564)	(0.168)	(0.152)	(0.162)	(0.164)	(0.0533)	(0.202)	(0.126)
Pacif × Mar	-0.199*	0.622***	-0.399**	0.221	0.794***	-0.225***	0.858***	0.00488
	(0.107)	(0.204)	(0.189)	(0.174)	(0.181)	(0.0742)	(0.248)	(0.155)
Pacif × Apr	-0.0515	0.644***	-0.310*	0.0493	0.920***	-0.236***	0.571**	0.134
	(0.0711)	(0.196)	(0.182)	(0.161)	(0.141)	(0.0625)	(0.215)	(0.142)
Pacif × May	0.0196	0.500**	-0.148	0.137	0.715***	-0.211***	0.734***	0.0469
	(0.0933)	(0.213)	(0.181)	(0.172)	(0.172)	(0.0658)	(0.196)	(0.198)
Pacif × Jun	-0.148***	0.443**	-0.397*	-0.173	0.672***	-0.161**	0.443**	-0.138
	(0.0464)	(0.183)	(0.197)	(0.158)	(0.210)	(0.0653)	(0.190)	(0.144)
Pacif × Jul	-0.0244	0.490**	-0.407**	-0.0711	0.821***	-0.183**	0.571***	-0.201
	(0.0691)	(0.182)	(0.160)	(0.178)	(0.135)	(0.0678)	(0.188)	(0.171)
Pacif × Aug	-0.0241	0.395**	-0.339*	0.0488	0.769***	-0.162**	0.742***	-0.0820
	(0.0779)	(0.178)	(0.179)	(0.182)	(0.157)	(0.0763)	(0.173)	(0.153)
Pacif × Sep	0.0996	0.319**	-0.330**	0.185	0.564***	-0.160***	0.788***	-0.0416
	(0.0747)	(0.153)	(0.162)	(0.144)	(0.153)	(0.0537)	(0.185)	(0.135)
Pacif × Oct	0.0112	0.540**	-0.124	0.0302	0.849***	-0.0898**	0.652***	0.0204
	(0.0700)	(0.220)	(0.134)	(0.173)	(0.150)	(0.0401)	(0.181)	(0.145)
Pacif × Nov	-0.0198	0.346*	-0.413**	-0.196	0.257	-0.233***	0.385**	-0.356*
	(0.0548)	(0.176)	(0.184)	(0.156)	(0.197)	(0.0736)	(0.181)	(0.182)
Pacif × Dec	-0.111	0.570***	-0.491***	0.152	0.750***	-0.0883	0.281	-0.187
	(0.0720)	(0.162)	(0.167)	(0.143)	(0.159)	(0.0586)	(0.178)	(0.173)
Intervention	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218	4218	4218	4218

The table presents the treatment effects obtained from the estimation of equation (6) for different crime indicators as the outcome variable, one in each column. The treatment variable, *Pacified*, is decomposed according to the twelve months of a year. We do not correct for the reporting bias for Murder, Police Action and Police Killings. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

M Alternative Specifications

OLS Without Log-Transformation. First, we assume that $Crime^{C}_{i,t} = Crime^{C,R}_{i,t} + NR^{C}_{i,t}$, where $NR^{C}_{i,t}$ is the non-reported share of events of category C. Then, the OLS specification without a logarithmic transformation that we would like to estimate is written as follows:

$$Crime^{C,R}_{i,t} = \alpha Intervention_{i,t} + \beta Pacified_{i,t} + X'_{i,t}\theta + \nu_i + \gamma_t - NR^{C}_{i,t} + \varepsilon_{i,t} \tag{7}$$

As before, we do not observe $RR^{C}_{i,t}$. Therefore, we use a proxy variable, $Accident^{R}_{i,t}$ to recover its variations. Now, we assume that the non-reported share of event C can be additively separated into three components as follows: $NR^{C}_{i,t} = NR^{C}_{i} + NR^{C}_{t} + NR_{i,t}$, where NR^{C}_{i} and NR^{C}_{t} are simply absorbed by the inclusion of the UPP and time fixed-effects. Again, the main assumption is that $NR_{i,t}$, the time-varying part of $NR^{C}_{i,t}$, is the same for categories of events. Then, we assume the following relation:

$$Accident^{R}_{i,t} = X'_{i,t}\lambda + d_i + d_t - NR_{i,t} + u_{i,t}$$

Then, this expression can be plugged into the first equation, to obtain the following one:

$$Crime^{C,R}_{i,t} - Accident^{R}_{i,t} = \alpha Intervention_{i,t} + \beta Pacified_{i,t} + X'_{i,t}(\theta - \lambda) + (\nu_i - d_i) + (\gamma_t - d_t) + (\varepsilon_{i,t} - u_{i,t}) \tag{8}$$

Estimating that equation provides an unbiased value of β. It corresponds to the solution that is presented in the main body of the paper to account for the endogeneity of the unobserved reporting rate.

It would also be possible to identify the value of the increase in the unobserved reporting rate. It involves presuming that the negative relationship between the time-varying part of the non-reported share of events and the treatment can be written as follows:

$$NR_{i,t} = -\delta_1 Intervention_{i,t} - \delta_2 Pacified_{i,t} + X'_{i,t}\omega + c_i + c_t + e_{i,t} \quad \text{with } \delta_1 \geq 0 \text{ and } \delta_2 \geq 0$$

By substituting this expression of $NR_{i,t}$ into the Accident equation, the following can be directly obtained:

$$Accident^R_{i,t} = \delta_1 Intervention_{i,t} + \delta_2 Pacified_{i,t} + X'_{i,t}(\lambda - \omega) + (d_i - c_i) + (d_t - c_t) + (u_{i,t} - e_{i,t})$$

Additionally, we could simultaneously estimate the biased value of α and β as well as the value of δ_1 and δ_2 in a SURE system to recover the unbiased value of α and β, as before.

Here, we present only the results from the main OLS specifications with and without the correction for the reporting bias in Table M.1. The results are very similar to that obtained with log-regressions.

Table M.1: Results obtained from OLS in level

	Panel A. Without correction of unobserved reporting rate				
	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.00000707***	0.000299***	-0.0000380	0.0000733***	0.00000397
	(0.00000238)	(0.0000440)	(0.0000356)	(0.0000251)	(0.00000330)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	0.000239***	-0.0000166***	0.000229***	0.0000151***	0.000999***
	(0.0000646)	(0.00000328)	(0.0000292)	(0.00000400)	(0.000104)
	Panel B. With correction of unobserved reporting rate				
	Murder	Assault	Robbery	Theft	Extortion
Pacified	N.R.	0.000269***	-0.0000674	0.0000439*	-0.0000255**
		(0.0000432)	(0.0000406)	(0.0000223)	(0.00000941)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	N.R.	N.R.	0.000199***	-0.0000143	0.000970***
			(0.0000324)	(0.00000868)	(0.000102)
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (7) in Panel A and of equation (8) in Panel B for different crime indicators as the outcome variable, one in each column. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP$\times$month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. N.R. stands for Not Relevant as the correction of the reporting bias is not relevant for some crime indicators. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Poisson Regressions. It is also possible to estimate the treatment effect of the pacification policy with Poisson regressions. First, let us assume that the number of crimes, denoted $crime_{i,t}$ with a small c, takes the following form:

$$crime_{i,t} = \exp(\alpha Intervention_{i,t} + \beta Pacified_{i,t} + X'_{i,t}\mu + \nu_i + \gamma_t) \times population_i \tag{9}$$

However, we are only able to observe the reported number of crimes such that we have the following:

$$\begin{aligned}
crime^R_{i,t} &= crime_{i,t} \times RR^C_{i,t} \\
&= \exp(\alpha Intervention_{i,t} + \beta Pacified_{i,t} + X'_{i,t}\mu + \nu_i + \gamma_t) \times population_i \times RR^C_{i,t} \\
&= \exp(\alpha Intervention_{i,t} + \beta Pacified_{i,t} + X'_{i,t}\mu + \nu'_i + \gamma'_t) \times population_i \times RR_{i,t}
\end{aligned}$$

As before, we do not observe the reporting rate; therefore, we use the number of accidents as a proxy variable. Assuming that the time-varying part of the reporting rate is the same for all categories of events, we have the following:

$$\begin{aligned}
accident^R_{i,t} &= \exp(X'_{i,t}\omega + d_i + d_t) \times population_i \times RR^A_{i,t} \\
&= \exp(X'_{i,t}\omega + d'_i + d'_t) \times population_i \times RR_{i,t}
\end{aligned}$$

Then, by substituting the expression of $RR_{i,t}$ into the main equation, we obtain the following:

$$crime^R_{i,t} = \exp\left(\alpha Intervention_{i,t} + \beta Pacified_{i,t} + X'_{i,t}(\mu - \omega) + (\nu'_i - d'_i) + (\gamma'_t - d'_t)\right) \times accident^R_{i,t} \tag{10}$$

which corresponds to the solution proposed in this paper to correct the endogeneity of the unobserved reporting rate. Identifying the value of the increase in the reporting rate is much more difficult to obtain as it would need to compute the value of the bias coming from the omission of a relevant explanatory variable in a Poisson regression model, which is beyond the scope of this paper. The number of accidents is introduced in the specification as an exposure variable. In general, an exposure variable A appears inside a log function in the log-likelihood of the Poisson regression model, and written program of statistical software usually maximizes the log-likelihood. Here, the number of accidents contains zeros so that the log-likelihood is undefined and statistical software, such as Stata, cannot estimate this specification.[5] Therefore, we have simply added a constant equal to 0.5 to all

[5] A solution would be to manually program the likelihood function and to maximize it, because the likeli-

accident observations so that the log-likelihood is defined. This procedure may introduce a bias in the value of the estimated β.

We present the results from Poisson specification with and without the correction for the reporting bias in Table M.2. The estimated coefficients are exponentiated, so their interpretation is straightforward, *i.e.*, they have a rate ratio corresponding to a one unit increase in the treatment variable. For instance, a coefficient equal to 1.3 (0.7) implies that the expected value of a given crime increases (decreases) by 30% following the treatment. Again, the results are very similar to that obtained with log-regressions, which confirms the robustness of the analysis. Nevertheless, the magnitude of some effects are substantially higher in absolute value for some crime indicators.

In a robustness check, we have deleted the observations with zero accidents and the results obtained in this case are very similar to those obtained when we add a constant equal to 0.5 to all accident observations. We can drop the observations with zero accidents without biasing the estimate of the β coefficient because the occurrence of an accident can be assumed to be random and independent from the realization of any crime.

As a final robustness check, we use Negative Binomial regressions to relax the assumption that the variance of the dependent variable is equal to the mean. The results are presented in Table M.3, they are very similar to those obtained from Poisson regressions and that are presented in Table M.2.

First difference estimator. The main estimated effects presented in Table 3 are obtained using a fixed effects (within) estimator that relies on the strong exogeneity condition. Therefore, we also estimate the pacification effect using a first difference estimator, which is well known for being less efficient but that needs a weaker condition than the strong exogeneity condition. The first difference estimators are presented in Table M.4, which also reports fixed effects estimators for ease of comparison. Although the standard errors are much higher with the first difference estimators, the point estimates are consistent between the first difference and fixed effect estimators, which lends credit to the strong exogeneity assumption.

hood function of the Poisson regression does involve any log function. However, it is often numerically more difficult to maximize a likelihood function rather than a log-likelihood function.

Table M.2: Results obtained from Poisson regressions

Panel A. Without correction of the reporting bias					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	0.605***	2.119***	0.804***	1.263***	1.346
	(0.0868)	(0.205)	(0.0440)	(0.0721)	(0.296)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	1.867***	0.134***	2.256***	1.567**	1.662***
	(0.277)	(0.0425)	(0.218)	(0.298)	(0.0958)
Panel B. With correction of the reporting bias					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	N.R.	1.698***	0.679***	1.013	1.079
		(0.197)	(0.0387)	(0.0788)	(0.272)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	N.R.	N.R.	1.848***	1.250	1.356***
			(0.246)	(0.253)	(0.109)
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (9) in Panel A and of equation (10) in Panel B for different crime indicators as the outcome variable, one in each column. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. Exponentiated coefficients. N.R. stands for Not Relevant as the correction of the reporting bias is not relevant for some crime indicators. Standard errors clustered by UPP in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table M.3: Results obtained from Negative Binomial regressions

Panel A. Without correction of the reporting bias					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	0.637***	2.081***	0.809***	1.283***	1.346
	(0.0913)	(0.200)	(0.0535)	(0.0778)	(0.296)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	1.932***	0.127***	2.269***	1.658***	1.719***
	(0.241)	(0.0421)	(0.225)	(0.303)	(0.101)
Panel B. With correction of the reporting bias					
	Murder	Assault	Robbery	Theft	Extortion
Pacified	N.R.	1.794***	0.619***	1.097	1.079
		(0.202)	(0.0643)	(0.0905)	(0.272)
	Police Action	Police Killings	Threat	Rape	Total Events
Pacified	N.R.	N.R.	2.098***	1.333	1.522***
			(0.279)	(0.245)	(0.114)
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	4218	4218	4218	4218	4218

Notes: The table presents the treatment effects obtained from the estimation of equation (9) in Panel A and of equation (10) using Negative Binomial regressions for different crime indicators as the outcome variable, one in each column.. All the regressions take into account the intervention period and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP×month, and the sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016. Exponentiated coefficients. N.R. stands for Not Relevant as the correction of the reporting bias is not relevant for some crime indicators. Standard errors clustered by UPP in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table M.4: First difference and fixed effect estimators

	Murder	Assault	Robbery	Theft	Extortion
Δ Pacified	-0.00727	0.694**	-0.371	-0.263	-0.158
	(0.0913)	(0.280)	(0.245)	(0.256)	(0.208)
Bias correction	No	Yes	Yes	Yes	Yes
Δ Intervention	Yes	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threat	Rape	Total Event
Δ Pacified	0.474*	-0.171	0.599**	-0.0331	0.274
	(0.262)	(0.106)	(0.243)	(0.220)	(0.201)
Bias correction	No	No	Yes	Yes	Yes
Δ Intervention	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	4181	4181	4181	4181	4181

Notes: Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The table presents the treatment effects obtained from the estimation of equation (6) using a first difference estimator for different crime indicators as the outcome variable, one in each column. We do not correct for the reporting bias for Murder, Police Action and Police Killings. All the regressions include time fixed effects. The sample includes observations for all UPP located in Rio de Janeiro between January 2007 and June 2016.

N Robustness Check Discarding the First Pacified Favelas

Batan and Cidade de Deus, the first favelas to be pacified, are located in a different area of the city and, therefore, might react differently to pacification than the other favelas. Table N.1 presents the results obtained by dropping Batan and Cidade de Deus, the first favelas to be pacified and that are located in a different area of the city than the others. The findings are also very close to the main results presented in Table 3.

Table N.1: Estimates obtained by dropping Batan and Cidade de Deus

	Murder	Assault	Robbery	Theft	Extortion
Pacified	-0.0391	0.469***	-0.340***	0.0256	-0.213***
	(0.0235)	(0.109)	(0.0955)	(0.0803)	(0.0610)
Bias correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threats	Rape	Total Events
Pacified	0.753***	-0.140***	0.550***	-0.0877	0.373***
	(0.141)	(0.0322)	(0.117)	(0.0622)	(0.0892)
Bias correction	No	No	Yes	Yes	Yes
Intervention	Yes	Yes	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations	3990	3990	3990	3990	3990

Notes: The table presents the treatment effects obtained from the estimation of equation (6) for different crime indicators as the outcome variable, one in each column. We do not correct for the reporting bias for Murder, Police Action and Police Killings. All the regressions take into account the intervention date and include UPP fixed effects, month fixed effects, and linear time trends specific to each UPP. The observation unit is the UPP$\times$month, and the sample includes observations for all UPP located in Rio de Janeiro, but excluding Batan and Cidade de Deus, between January 2007 and June 2016. Standard errors clustered by UPP are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

O ROBUSTNESS CHECKS ON THE GANG GOVERNANCE EFFECT

Table O.1: Results on the favelas controlled by ADA or CV before pacification

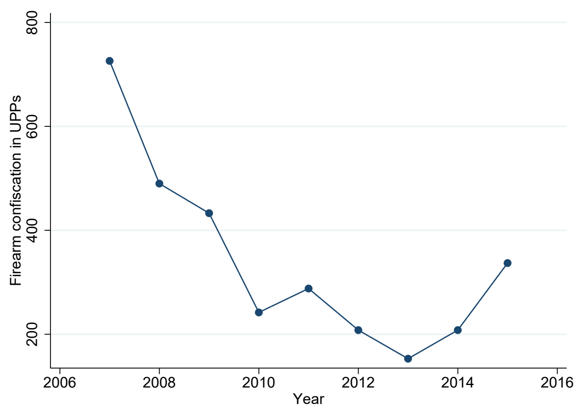
Panel A. Favelas controlled by ADA before pacification					
	Murder	Assault	Robbery	Theft	Extortion
Intervention	0.144	1.286**	0.245	1.039**	-0.0597
	(0.124)	(0.434)	(0.519)	(0.374)	(0.128)
Pacified	-0.0301	1.018**	-0.298	0.590**	-0.304
	(0.0457)	(0.250)	(0.176)	(0.183)	(0.152)
Bias correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threats	Rape	Total Events
Intervention	1.864***	0.102	1.265**	0.217	0.938**
	(0.196)	(0.127)	(0.387)	(0.105)	(0.272)
Pacified	1.771***	-0.0177	0.837	-0.0553	0.585**
	(0.185)	(0.0706)	(0.441)	(0.158)	(0.180)
Bias correction	No	No	Yes	Yes	Yes
Panel B. Favelas controlled by CV before pacification					
	Murder	Assault	Robbery	Theft	Extortion
Intervention	-0.0417	0.419***	-0.0180	0.235**	0.0243
	(0.0520)	(0.138)	(0.161)	(0.0992)	(0.0745)
Pacified	-0.0761	0.385***	-0.337***	-0.0622	-0.231***
	(0.0465)	(0.123)	(0.0973)	(0.0817)	(0.0591)
Bias correction	No	Yes	Yes	Yes	Yes
	Police Action	Police Killings	Threats	Rape	Total Events
Intervention	0.722***	-0.112**	0.594***	0.134	0.507***
	(0.152)	(0.0474)	(0.128)	(0.0828)	(0.121)
Pacified	0.594***	-0.201***	0.460***	-0.115	0.300***
	(0.147)	(0.0500)	(0.122)	(0.0678)	(0.104)
Bias correction	No	No	Yes	Yes	Yes
UPP fixed effects	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
UPP linear time trends	Yes	Yes	Yes	Yes	Yes
Observations (Panel A)	570	570	570	570	570
Observations (Panel B)	3192	3192	3192	3192	3192

Notes: Results of regressions obtained by estimating the effect of treatment only with the favelas controlled by ADA or by CV before pacification. Standard errors clustered by UPP in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

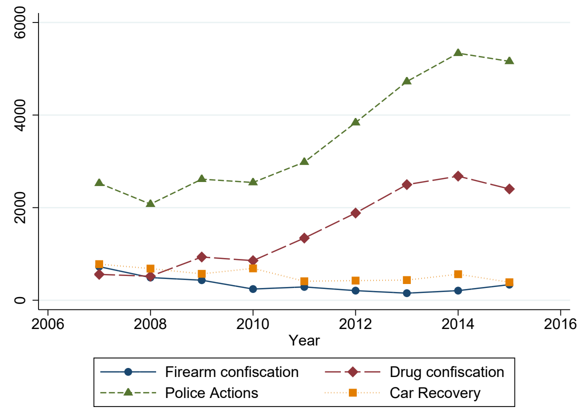
P Graphical Evidence on Firearm Confiscation

Figure P.1: Police action in UPPs

(a) Firearm confiscation

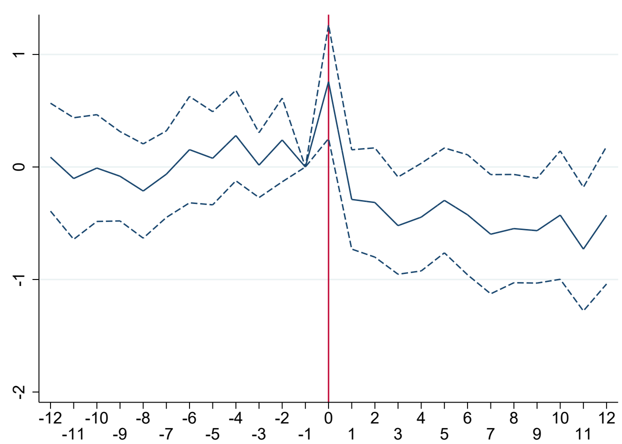


(b) Police action



Notes: Figures plot the annual police activity in Rio de Janeiro in areas that were covered by UPPs at the end of the study period.

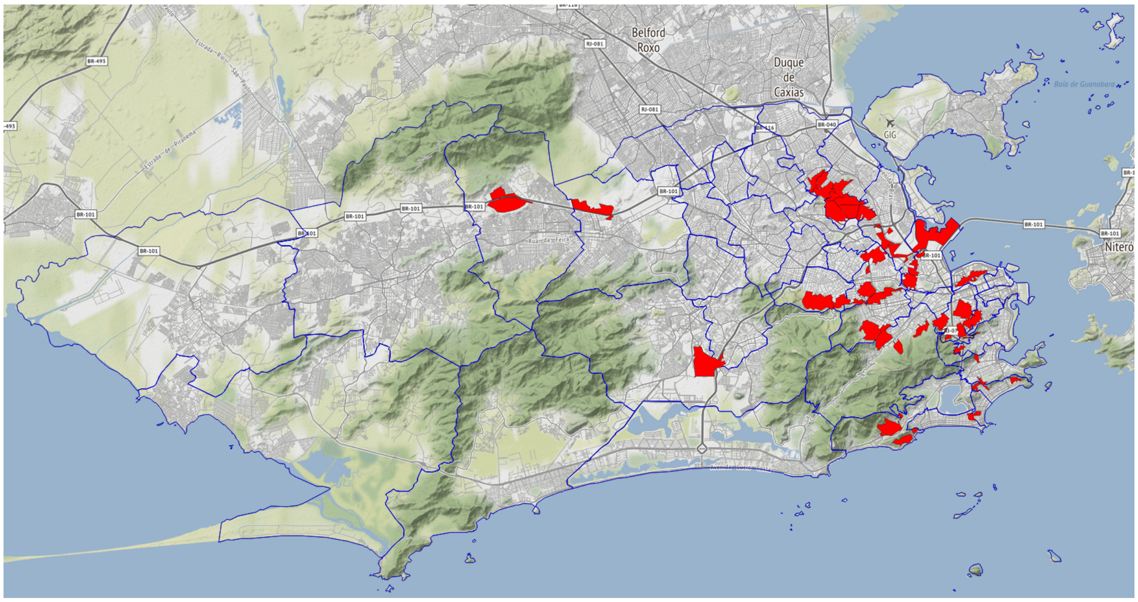
Figure P.2: Dynamic treatment effects on firearm confiscation



Notes: The figure plots the quarterly rate for the number of firearms confiscated as a function of time since BOPE intervention. The solid lines correspond to the values of π_k $(k < 0)$ and τ_k $(k \geq 0)$, as a function of k, obtained from the estimation of equation (7) on the sample of all UPPs for which $k \in [-12, +12]$. Standard errors are clustered at the UPP level and dashed lines represent the 95% confidence interval.

Q Districts and UPPs

Figure Q.1: Mapping between districts and UPPs



Notes: Districts of Rio de Janeiro are delimited by the blue lines and all UPPs installed within Rio de Janeiro at the end of the policy are drawn in red.

References

Ayres, I. and S. D. Levitt (1998): "Measuring Positive Externalities from Unobservable Victim Precaution: An Empirical Analysis of Lojack," *The Quarterly Journal of Economics*, 113, 43–77.

Bellégo, C., D. Benatia, and L.-D. Pape (2022): "Dealing with the Logs and Zeros in Regression Models," *CREST Working Paper*.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 90, 414–427.

de Chaisemartin, C. and X. D'Haultfœuille (2020): "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–96.

DiPasquale, D. and E. L. Glaeser (1999): "Incentives and Social Capital: Are Homeowners Better Citizens?" *Journal of Urban Economics*, 45, 354–384.

DRACA, M., S. MACHIN, AND R. WITT (2011): "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks," *American Economic Review*, 101, 2157–81.

EHRLICH, I. (1973): "Participation in Illegitimate Activites: A Theoretical and Empirical Investigation," *Journal of Political Economy*, 81, 521–565.

FISHER, R. A. (1935): *The Design of Experiments*, Macmillan.

GLAESER, E. L. AND B. SACERDOTE (1999): "Why Is There More Crime in Cities?" *Journal of Political Economy*, 107, 225–258.

JARVIS, A., H. I. REUTER, A. NELSON, AND E. GUEVARA (2008): Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT) (accessed July 2018) , `https://srtm.csi.cgiar.org`.

LEVITT, S. D. (1998): "Juvenile Crime and Punishment," *Journal of Political Economy*, 106, 1156–1185.

LOCHNER, L. AND E. MORETTI (2004): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *The American Economic Review*, 94, 155–189.

POSNER, R. A. (2010): "From the New Institutional Economics to Organization Economics: With Applications to Corporate Governance, Government agencies, and Legal Institutions," *Journal of Institutional Economics*, 6, 1–37.