Did the Military Interventions in the Mexican Drug War Increase Violence?

Valeria Espinosa[1] and Donald B. Rubin[1]

Department of Statistics, Harvard University

Abstract

We analyze publicly available data to estimate the causal effect of military interventions on the homicide rates in certain problematic regions in Mexico. We use the Rubin Causal Model (Rubin 1974) to compare the post intervention homicide rate in each intervened region to the hypothetical homicide rate for that same year had the military intervention not taken place. Because the effect of a military intervention is not confined to the municipality subject to the intervention, a non-standard definition of units is necessary to measure plausibly the causal effect of the intervention under the standard no-interference assumption of SUTVA. Donor pools for each missing potential outcome under no intervention are created, thereby allowing for the estimation of unit-level causal effects. A multiple imputation approach accounts for the uncertainty about the missing potential outcomes.

Key words: causal inference, donor pools, Mexican Drug War.

1   Introduction

The Mexican presidency of Felipe Calderón (2006-2012) was characterized by its war against organized crime, raising many questions regarding security and violence. It is estimated that during this period, the war claimed 60,000 lives[1]. Mexico has 31 states and a Federal District, which are further partitioned into municipalities and delegations, respectively. Throughout this document, we will refer to all 32 federal entities as states, and all their political subdivisions as municipalities.

In 2011, two articles were published in a leading Mexican magazine, *Nexos*, on the effect of the military interventions on civilian homicides. Through visual comparisons, but no formal statistical analysis, Escalante (2011) explored the possibility that military interventions had increased the homicide rates in those states where the interventions took place. Later, Merino (2011) reached a similar conclusion using propensity score methods, a commonly used tool to help causal inference.

In both articles, the question of interest is whether the military interventions increased the homicide rate in those states where the interventions took place, beyond what those homicide rates would have been during the same post intervention period without the interventions. For example, consider the state of Chihuahua, where more than one military intervention occurred; both articles attempt to compare the post intervention homicide rate in Chihuahua to the counterfactual homicide rate in Chihuahua during the same time period had the state not experienced any military intervention. In the Rubin Causal Model, RCM, these homicide rates correspond to two potential outcomes for each state, one under active treatment (at least one military intervention), $HR_{state}(1)$, and one under control (no military intervention), $HR_{state}(0)$. The fundamental problem of causal inference is that only

---

[1]The Washington Post article can be found here: `http://www.washingtonpost.com/world/the_americas/calderon-finishes-his-six-year-drug-war-at-stalemate/2012/11/26/82c90a94-31eb-11e2-92f0-496af208bf23_storHR_1.html`

one of these potential outcomes is observed and therefore, implicitly or explicitly, the other missing potential outcome must be imputed to draw causal inferences. In this paper, we use the RCM and propensity scores to analyze this question for all "regions", defined to be collections of municipalities within a state, where interventions took place.

Propensity scores (Rosenbaum and Rubin 1983) are used in causal inference to create subsets of treated and control units with similar covariate values, for instance, here regions with similar characteristics. The propensity score is a function of the covariates with the property that units with the same value of the propensity score will have, in expectation, the same values on covariates used to define, or to estimate, the propensity score. When covariate values in a subset of treated and control units are similar, we say that there is *covariate balance* in that subset. Then, within such subsets, the observed outcomes of the control units can be used to impute the missing control outcomes for the treated units, and the observed values of outcomes for the treated units can be used to impute the missing values of treated outcomes for the control units. The main goal of propensity scores is to achieve covariate balance by design. For any particular dataset though, this balance should be checked to verify that we have created an appropriately balanced design.

Using propensity scores, Merino (2011) presents a pressumably causal inference analysis. However, he does not mention or assess balance, either analytically or visually, in any way. Also, his propensity score is a function of only one variable: the pre-intervention homicide rate. The definition of pre-intervention homicide rate is unclear for the control units because the military interventions took place at varying points in time. Moreover, the use of only one covariate first, makes the plausibility of unconfoundedness (see Section 3.2)- a critical assumption being made for causal inference- questionable, and second, there is no benefit when using propensity scores compared to matching on the single covariate.

Both Escalante (2011) and Merino (2011) focus on the state as the unit of analysis and both base their conclusions on averages across states. However, it is more appropriate

to treat the municipality as the unit of analysis because military operations tend to be directed at municipalities rather than entire states. Nonetheless, it is reasonable to assume that a military intervention focused on a particular municipality also has an effect on its neighboring municipalities. To capture the effect on the municipality that was directly intervened as well as its neighbors, we define the unit of analysis to be the municipality that received the intervention together with its immediate neighbors – we call this collection of municipalities a region. This definition of the unit of analysis addresses possible *interference* between municipalities, which can complicate a plausible causal analysis. To impute the missing non-intervention potential outcome for each treated region, we create donor pools of hypothetical control regions. These hypothetical control regions are generated by randomly selecting for each municipality in the treated region a control municipality from a set of acceptable matches of comparable quality. In addition to reporting estimates of the average effect across the regions, we also report estimates of the causal effect for each region, allowing the analyst to compare the estimated effectiveness of interventions in, for instance, Juárez and Acapulco.

In accordance with the RCM, this observational study can be viewed as a broken randomized experiment, broken because the assignment of the treatment (military intervention) was not random but occurred by some political process. As such, the distribution of covariates in the active treatment group (regions that were subject to military interventions) and the control group (those that did not) may in fact have unbalanced covariate distributions. The Mexican government had limited resources, and therefore military interventions were not sent to every municipality facing drug cartel related violence, but to those municipalities believed to be of highest priority by the government, and therefore are expected to differ from non-intervened municipalities on baseline covariates. The matching method attempts to find subsets of these municipalities with overlap in the distribution of key covariates.

In Section 2, we discuss the covariates, the data sources and the role that subject-matter

experts played in the design and analysis of our study. Section 3 focuses on design issues including the definition of estimands, key assumptions, and the assessment of covariate balance. Section 4 introduces the estimation procedure, and Section 5 outlines the results. Section 6 contains conclusions and future steps.

2   Data

We used publicly available data from three reliable sources: 1) the National Institute of Statistics and Geography (INEGI), 2) the Center of Research for Development (CIDAC), and 3) the Presidency website (CPWS). The data from the third source were available until the December 2012 presidential transition.

2.1   Covariates

There are 2456 municipalities in Mexico [2]. For each of these municipalities, we have covariates related to demographics, economics, location, education, health, politics and roads. We also considered as relevant three variables at the state level. All covariates are listed in Table 1, together with their sources; as shown there, INEGI was the source of most of the covariates. Escalante (2011) and Merino (2011) also used INEGI's homicide data[3].

Due to security concerns, certain covariates were not publicly available. Examples include smuggling routes, drug crop locations, and the specific cartel(s) present in each region. Graphical information about cartel location after 2006 can be found online in maps created by Stratfor[4], but we did not find any pre-intervention raw data. Coscia and Rios (2012) made

---

[2] http://www.inegi.org.mx/inegi/contenidos/espanol/prensa/Contenidos/capsulas/2002/geografica/municipios.asp

[3] Although Merino (2011) compared the results obtained with INEGI's data - based on death certificates - to those obtained with the National System for Public Security - based on preliminary investigations of public ministries. In a separate analysis, he also used the federal government database (PWS), which attempts to count the homicides associated with organized crime, but we used only pre intervention values as background covariates.

[4] A couple of their cartel related maps can be found in http://mexicanosalgrito.wordpress.

Type	Municipality Level	Source	State Level	Source
Demographics	Homicide Rate 2006*	INEGI	Homicide Rate 2006*	INEGI
	Indicator of Homicide Rate 2006 above the mean	INEGI		
	Population 2006	INEGI	Population 2005	INEGI
	Criminal Rivalry Death Count* (Dec 2006, Jan, Feb, Mar, Apr 2007)	CPWS		
Economics	Government Spending 2006	INEGI	GDP 2006	INEGI
Location	Latitude	INEGI		
	Longitude	INEGI		
Education	Average Years Of Schooling	INEGI		
	Proportion that can read and write	INEGI		
	Proportion that speaks an indigenous language	INEGI		
Health	Number of Doctors per medical unit	INEGI		
	Missingness indicator for the above	INEGI		
Politics	**Political Party in power at the end of 2006**	CIDAC		
Roads	Longitude of Roads	INEGI		
	Missingness Indicator of above	INEGI		

Table 1: Covariates used in the matching procedure. Entries in bold were exactly matched. * The municipality level homicide counts, total and criminal rivalry related, were transformed to homicide rates using the 2005 population information.

a serious attempt to track the cartels at the municipality level by web scraping Google News. We decided not to include these data because they are not publicly available. Obtaining these variables in a reliable and time-relevant form would, we believe, improve the quality of our study.

2.2   Treatment Indicator and Outcomes

The treatment indicator, commonly denoted W, here identifies whether a municipality received a military intervention ($W = 1$) or not ($W = 0$). The encoding of this variable was based on the military interventions listed in Escalante (2011), which was obtained as a result of a web search of press releases and "comunicados" of the Office the Secretary of Defense (SEDENA). A possible improvement of our study would be to conduct a web search of SEDENA press releases and Google News ourselves, to obtain a more comprehensive list of interventions. The true list is classified information.

com/2011/09/06/el-mapa-actualizado-de-los-carteles-de-las-drogas-septiembre-2011/   and
http://1.bp.blogspot.com/_DwzFG3eL8Do/TNh5MAu09KI/AAAAAAAAS4/pgUDlYgC5YU/s1600/5-17-10_
Mexican-drug-cartels-map_manufacturing_v5b.jpg for 2007 and 2010 respectively.

The homicide rate per 100,000 inhabitants is a standard measure of violence. For example, according to Wikipedia the homicide rate in Mexico in 2011 was 24, and the corresponding ones for Brazil, Colombia, Honduras and the United States are 21.8, 33, 91.6 and 4.2 respectively[5]. Nevertheless, this average measure can hide great annual variations within a country. For example, in 2011 the homicide rate in Juárez, México was 147.77, in Marceió, Brazil it was 135.26, in Cali, Colombia it was 77.90, in San Pedro Sula, Honduras it was 158.87, and in New Orleans, USA it was 57.88.[6] In this study, we attempt to assess whether military interventions increased violence measured by homicide rates per 100,000 inhabitants. Hence, our main outcome variable is the homicide rate one year after the military intervention. These data were also obtained at the municipality level from INEGI.

2.3 Subject matter experts

Rubin (2008) emphasizes the importance of subject matter knowledge in the outcome-free design of a causal study. We discussed the project with three experts in the field: Miguel Basañez and two graduate students at Harvard who have been working on these topics, Elisa de Anda and Viridiana Ríos. For example, they suggested including the covariate *political party* in office at the end of 2006, for each municipality. However, they were not otherwise actively involved in the design process.

3   Design

Randomized experiments are generally considered to be the gold standard for obtaining objective causal inferences. However, in this setting, a randomized experiment is not possible,

---

[5]The figures for Mexico, Brazil and Honduras were obtained from `http://www.usatoday.com/story/news/world/2013/07/31/mexico-murder-rate/2606229/`, the others can be found in Wikipedia, the US in `http://blogs.justice.gov/main/archives/1765`...the 2010 data in `http://www.refworld.org/docid/50bf2c812.html` for Brazil and `http://www.unodc.org/documents/data-and-analysis/statistics/Homicide/Globa_study_on_homicide_2011_web.pdf`.

[6]These numbers were obtained from Wikipedia.

and an observational study is the only way to assess the question of interest. Following Rubin (2007, 2008), we attempt to approximate an experimental approach as closely as possible by separating the design phase from the analysis phase. The key distinction is that in the design stage all outcome data are absent. Therefore, in the design of this study we only used data from 2006 or earlier because all interventions occurred after 2006, and any post 2006 variables are possibly affected by the intervention or its absence.

To frame the question of interest one must define the treatment, units of analysis and estimand. Here, we define the treatment to be a military intervention, specifically a confrontation between army and organized crime that resulted in at least three civilian deaths (where civilian could refer to a member of a cartel) based on the list of municipalities where confrontations occurred given in Escalante (2011). Note that this definition is different from "sending military forces" to a municipality, which would be the ideal definition of treatment for this study. However, a comprehensive list of such events is classified, so we do not have access to it. The use of the list given in Escalante (2011) is appealing because we make the same distinction between intervened and not-intervened municipalities that the original motivating *Nexos* paper did (it is unclear what definition Merino (2011) used).

3.1 Treated Units and Potential Outcomes

We define the treated units (i.e., the treated regions) as the group of contiguous municipalities where at least one municipality received a direct military intervention. The definition of treatment for a region is receiving at least one military intervention between 2007 and 2010, and the time of treatment is the first time a municipality in the region received an intervention. Let $HR_i^t(W_i)$ denote the homicide rate for treated region i at the *end* of year t under treatment level W_i (see Figure 2), where $W_i = 0, 1$ ($W_i = 1$ is observed for treated regions), and let $+$ generically denote "one year after" the military intervention. We are interested in comparing the observed homicide rate in region i one year after intervention, $HR_i^+(1)$,

7

to the homicide rate in the same year if it had not received the intervention, $HR_i^+(0)$. In other words, the estimands of interest are $HR_i^+(1) - HR_i^+(0)$, at the region level, and the average across all treated regions $\tau = \frac{\sum_{i=1}^{N} HR_i^+(1) - HR_i^+(0)}{N}$, where N denotes the total number of treated regions.

3.2   Key assumptions: SUTVA and Unconfoundedness

In this study we make two key assumptions commonly made in this framework: *the stable-unit treatment value assumption*, SUTVA - which is implicitly made in the notation introduced in the previous section, and unconfoundedness. SUTVA is fundamental in this formulation of the problem, and unconfoundedness is crucial to understand the role of relevant background covariates.

SUTVA comprises two parts: no interference between units and no hidden versions of treatment. *No interference* means that each unit's treatment assignment has no effect on the outcomes of other units, and *no hidden versions of treatment* means that there is an unambiguous definition of treatment, thereby allowing each unit to be clearly identified as intervened or not. That is, each unit's potential outcomes are functions of the unit label, i, and the treatment the unit received, W_i. For the plausibility of this assumption, it is important to think carefully about the definition of treatment and the units that receive it, as we discussed earlier. For instance, it is reasonable to assume that the effect of a military intervention on a particular *municipality* would spill over to neighboring municipalities; hence, we grouped municipalities that received an intervention with their immediate neighbors to form treated regions for which "no interference" is more plausible. This definition may force treated regions to comprise more than one municipality that is directly intervened. A more ambitious approach could consider this as a case of a treatment with multiple levels. However, in this study, we aim to provide a simple analytical approach by providing a definition of a binary treatment that is broad enough to classify all treated regions into
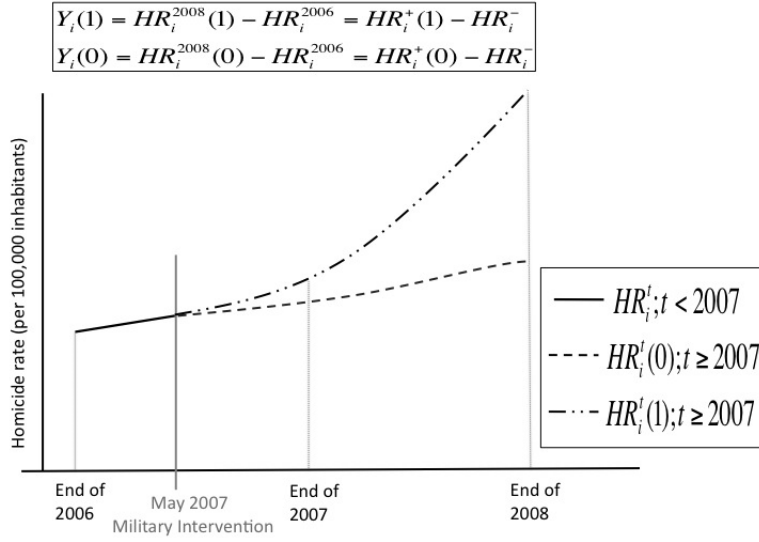
$$Y_i(1) = HR_i^{2008}(1) - HR_i^{2006} = HR_i^+(1) - HR_i^-$$
$$Y_i(0) = HR_i^{2008}(0) - HR_i^{2006} = HR_i^+(0) - HR_i^-$$

Table 2: **Sketch of the definition of the potential outcomes.** For each region i, HR_i denotes the homicide rate and the Y_i denotes the post - pre intervention difference in homicide rate (i.e., the homicide rate one year after the military intervention minus the homicide rate one year before the intervention). There are three types of lines. The solid one corresponds to the pre-intervention homicide rates, and the upper and lower dashed lines correspond to the homicide rates with and without the military intervention, respectively.

Region	$HR_i^+(0)$	$HR_i^+(1)$	$Y_i(0)$	$Y_i(1)$	X_i
Tijuana	$HR_1^+(0)$	$HR_1^+(1)$	$Y_1(0)$	$Y_1(1)$	X_1
Nogales	$HR_2^+(0)$	$HR_2^+(1)$	$Y_2(0)$	$Y_2(1)$	X_2
Madera	$HR_3^+(0)$	$HR_3^+(1)$	$Y_3(0)$	$Y_3(1)$	X_3
Juárez	$HR_4^+(0)$	$HR_4^+(1)$	$Y_4(0)$	$Y_4(1)$	X_4
Pánuco	$HR_5^+(0)$	$HR_5^+(1)$	$Y_5(0)$	$Y_5(1)$	X_5
Reynosa	$HR_6^+(0)$	$HR_6^+(1)$	$Y_6(0)$	$Y_6(1)$	X_6
Bustamante	$HR_7^+(0)$	$HR_7^+(1)$	$Y_7(0)$	$Y_7(1)$	X_7
Guadalupe	$HR_8^+(0)$	$HR_8^+(1)$	$Y_8(0)$	$Y_8(1)$	X_8
Villa de Cos	$HR_9^+(0)$	$HR_9^+(1)$	$Y_9(0)$	$Y_9(1)$	X_9
Teúl	$HR_{10}^+(0)$	$HR_{10}^+(1)$	$Y_{10}(0)$	$Y_{10}(1)$	X_{10}
Rincón de Romos	$HR_{11}^+(0)$	$HR_{11}^+(1)$	$Y_{11}(0)$	$Y_{11}(1)$	X_{11}
Sinaloa	$HR_{12}^+(0)$	$HR_{12}^+(1)$	$Y_{12}(0)$	$Y_{12}(1)$	X_{12}
Tepic	$HR_{13}^+(0)$	$HR_{13}^+(1)$	$Y_{13}(0)$	$Y_{13}(1)$	X_{13}
La Piedad	$HR_{14}^+(0)$	$HR_{14}^+(1)$	$Y_{14}(0)$	$Y_{14}(1)$	X_{14}
Celaya	$HR_{15}^+(0)$	$HR_{15}^+(1)$	$Y_{15}(0)$	$Y_{15}(1)$	X_{15}
Apatzingán	$HR_{16}^+(0)$	$HR_{16}^+(1)$	$Y_{16}(0)$	$Y_{16}(1)$	X_{16}
Coahuayana	$HR_{17}^+(0)$	$HR_{17}^+(1)$	$Y_{17}(0)$	$Y_{17}(1)$	X_{17}
Acapulco	$HR_{18}^+(0)$	$HR_{18}^+(1)$	$Y_{18}(0)$	$Y_{18}(1)$	X_{18}

Table 3: **Science Table for the treated regions.** This table symbolically contains the potential outcomes and pre-intervention covariates for all regions. $HR_i^+(1)$ denotes the homicide rate one year after receiving a military intervention, and $HR_i^+(0)$ denotes the homicide rate in the same year not receiving a military intervention. $Y_i(1)$ and $Y_i(0)$ represent the one year post - one year pre intervention differences in homicide (observed) respectively and what this difference would have been without the intervention. X_i denotes the background covariates of region i (i.e., the list of all the municipality level covariates described in Table 1).

the same treatment level. The treated regions are formally defined as contiguous, but for reasons discussed in Section 3.4, the hypothetical control regions we used are allowed to be formed by discontiguous municipalities that were not intervened.

To understand the unconfoundedness assumption, let $\mathbf{X}$ denote the matrix of covariates, where each row corresponds to a region and each column to a variable. Similarly, let $\mathbf{Y}$ denote the $N \times 2$ matrix where each row corresponds to the two potential outcomes for each region. Unconfoundedness means that, given all observed covariates $\mathbf{X}$, treatment assignment is functionally independent of $\mathbf{Y}$, the potential outcomes. If unconfoundedness holds, a formal causal interpretation of the results is possible. If unconfoundedness does not hold, a causal interpretation is no longer valid, although a principled conditional association interpretation can still be made. As we have previously mentioned, there are unavailable covariates that we believe might be important to improve the plausibility of unconfoundedness in our study. Because these variables were also missing in other assessments of the causal question of interest, we need to be cautious when making formal causal interpretations.

3.3   Estimands

Table 3, which is often called the SCIENCE in the RCM framework, shows all the quantities we need to know to *calculate* the TRUE value of the estimands. However, we did not observe any of the outcomes for these regions when not receiving the military intervention. Hence, the $HR_i^+(0)$ column of Table 3 is completely missing and each of these quantities needs to be estimated, i.e., imputed.

Furthermore, let N_i denote the number of municipalities that form region i, and $Pop_{ij}^+(W_i)$ and $H_{ij}^+(W_i)$ the population size and number of homicides of the j-th municipality in the i-th region one year after receiving treatment level W_i. Also, letting $Pop_i^+(W_i) = \sum_j^{N_i} Pop_{ij}^+(W_i)$ denote the total population one year after i receives treatment W_i, the potential outcomes

can be expressed as

$$HR_i^+(1) = \frac{\sum_{j=1}^{N_i} H_{ij}^+(1)}{\sum_{j=1}^{N_i} Pop_{ij}^+(1)}, \text{ and } HR_i^+(0) = \frac{\sum_{j=1}^{N_i} H_{ij}^+(0)}{\sum_{j=1}^{N_i} Pop_{ij}^+(0)}.$$

The use of covariates can be beneficial for precise estimation purposes. Subtracting from the outcome a pre-intervention variable that is a lagged version of the outcome (in this case, the homicide rate the year before the intervention took place) will usually increase the efficiency of the estimator without modifying the estimand. See Section 4 for details of how this was used in our estimation procedure.

3.4 Matching Method

The goal of our matching method is to create hypothetical control regions that are similar to the treated ones on background covariates. First, using a combination of exact and propensity score matching, from a pool of 2208 municipalities, we found two control municipalities of comparable quality to match to each of the 248 municipalities forming the treated regions.

We exactly matched on three variables: political party in office at municipality level in 2006, the indicator of whether the 2006 municipality's homicide rate was above the national mean, and the missingness indicator of the number of doctors per medical unit in 2005 (1 if missing and 0 otherwise). The second and final step was to select two control municipalities to match each municipality in every treated region. For each municipality within a treated region, the aforementioned exact matching defined a set of control municipalities out of which two matches were selected based on proximity on their estimated propensity score to that of the target municipality in the treated region. When estimating the propensity score, in addition to main effects that were not exactly matched, we used estimated second order effects involving the 2006 homicide rate at the municipality and some transformations of

the 2006 homicide rates at the state and municipality levels[7]. The final choice of propensity score was based on an assessment of the quality of balance achieved, where details are given in the next section. For every treated region, a donor pool of control regions was created and the eventual matches were created by randomly drawing one of the control matches associated with each of the municipalities forming the treated region (described in Section 4).

3.5   Balance Assessment

As in any observational study, balance on key covariates should be assessed before any examination of outcome data. When reasonable balance has been achieved, the design stage of the study ends and the analysis stage begins.

The assessment of balance was made at the municipality level because the matching was done at that level, which was the highest level of resolution in the raw data. "Love plots" (Ahmed et al. 2006) and histograms are useful tools to evaluate balance. In this context, the Love plots compare the difference in the covariate means between the municipalities in treated regions and the control regions before matching (i.e., all the municipalities in Mexico that were not in a treated region) and after matching (i.e., the subset of municipalities that were matched to the those municipalities forming treated regions). The comparison for binary covariates is based on the simple difference in means, whereas the comparison of continuous covariates is based on the standardized difference in means using the estimates of the variances calculated before matching. The smaller these differences are, the better, which in the Love plot is represented by the proximity of the symbols to the vertical line at zero, which denotes no difference in the average values of the covariates between municipalities in the treated regions and the set of its controls.

---

[7] We considered all main effects, second order terms involving Homicide Rate 2006 at Municipality level, log(Homicide Rate 2006 +1), log(State Homicide Rate 2006 +1), and State Homicide Rate 2006^2.

Figure 1 displays the Love plots for our study, where the pre-matching values are represented by gray squares and the post-matching values by black dots. These plots show a clear improvement post-matching for every covariate, except for the 2006 State GDP (where the balance worsens) and the criminal rivalry deaths (where the improvement is minor) . This conclusion is revealed by the fact that the post-matching (black dots) points are closer to the vertical line than the pre-matching points (grey squares). In particular, there is a big improvement on the variables involving pre-intervention homicide rates. Subfigure 1(a) shows that the difference in means for the State GDP in 2006 increased after matching, as mentioned previously. The balance on the monthly *criminal rivalry* related homicide rates before May 2007 (time of first intervention) did not improve as much as the balance for annual homicide rates; in particular, the improvement for the January 2007 value is only of 0.037, which is almost imperceptible in Figure 1(c). Nevertheless, we chose to focus on balancing the 2006 homicide rate reported by INEGI because those are the official data, and the reliability of the criminal rivalry death count is unclear.

Figure 2 shows the propensity score distribution by region, where the control region distributions (grey histograms) consist of twice as many municipalities as the intervened regions (black histograms). In some cases the number of municipalities within a treated region is very small (e.g., Tijuana consists of 4). The balance for the municipalities in the Tijuana, Apatzingán and Nogales regions is questionable because of the lack of overlap in the region of high propensity scores.
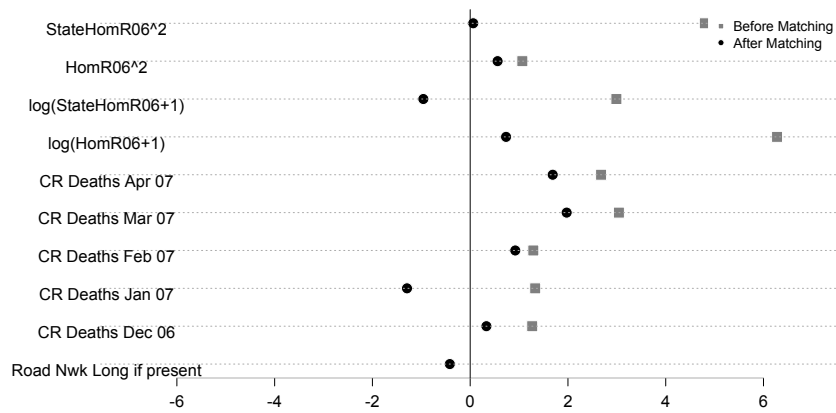
4 Estimation

The use of pre-treatment covariates to predict the observed outcomes can help increase the efficiency of the estimation. However, in this case a careful modeling of the outcome needs to take into account the interference between neighboring municipalities in the treated regions.
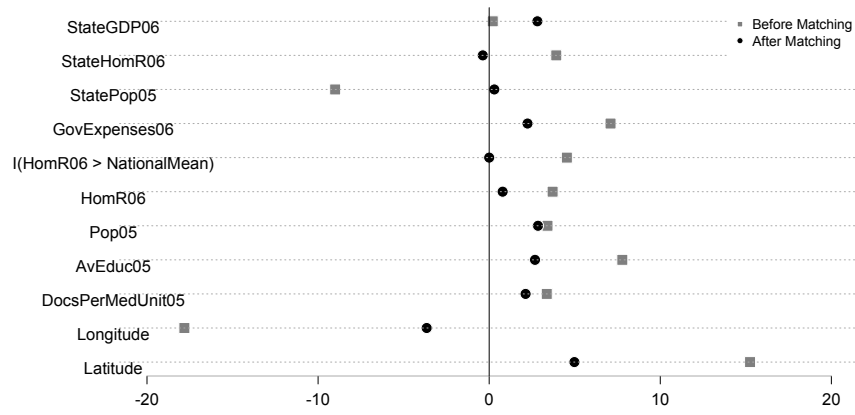
(a) Continuous Covariates
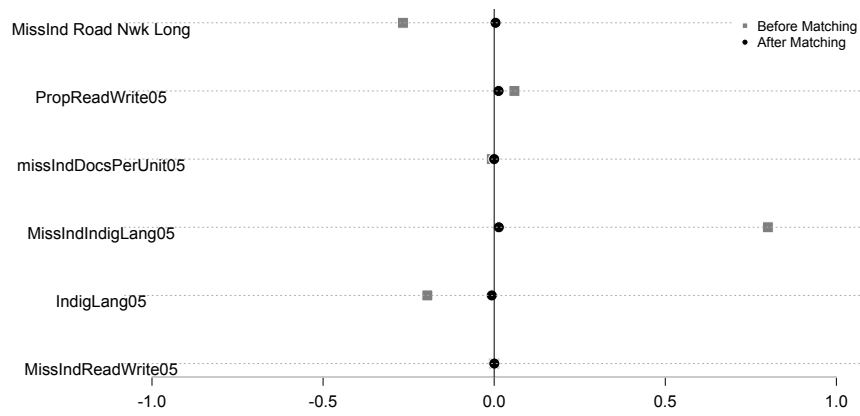
(b) Proportions and Binary Covariates

(c) Additional covariates

Figure 1: Love plots comparing the difference in means in covariates pre and post matching. The "CR deaths" terms refer to criminal rivalry related deaths.

Propensity Score Distributions for Regions 1 to 9

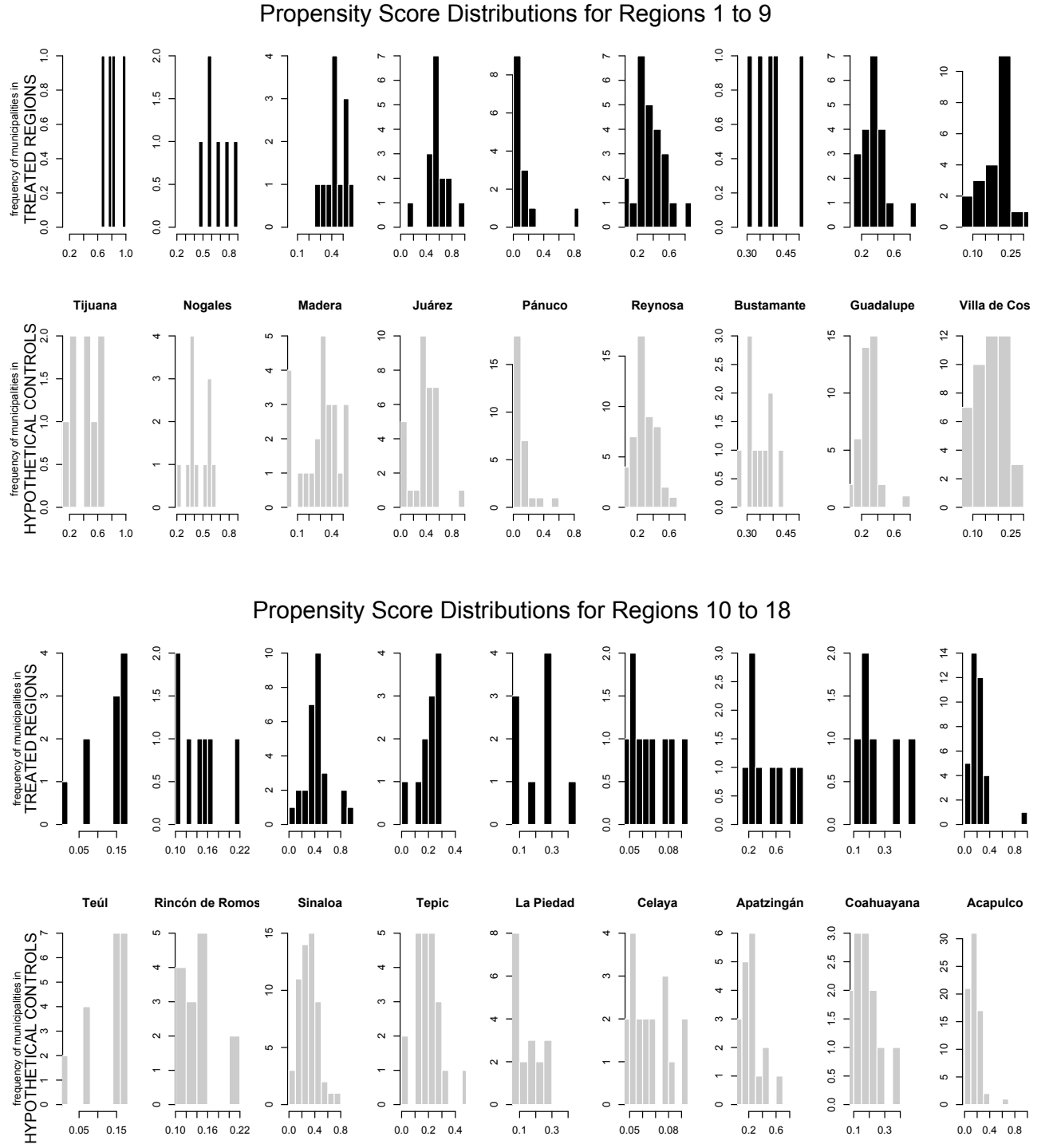

Propensity Score Distributions for Regions 10 to 18

Figure 2: **Propensity score histogram by Region**. Two histograms are displayed (vertically) for each region. Intervened regions (black) are on top, and control ones (gray) are on the bottom.

Such modeling is contrary to the goal of this paper, which is to use a *simple* analysis that is more statistically valid that the ones previously done. Therefore, and with efficiency in mind, we only use one predictive pre-treatment covariate to transform the observed outcome in to what is commonly referred to as *gain scores*. That is, we subtract a lagged version of the outcomes (i.e., measured prior to assignment to treatment or control), from the original potential outcomes to define the post - pre intervention differences in homicide rates, and then carry out the analysis on these transformed outcomes. This modification has an effect on the estimation but not on the estimand. In this case, let HR_i^- denote the homicide rate for region i one year before treatment is received, therefore being the same when $W_i = 1$ or 0. We transformed the outcomes and defined the modified potential outcomes for region i as the post - pre difference observed in the homicide rate in region i between one year after receiving a military intervention and the rate one year before receiving the intervention, $Y_i(1) = HR_i^+(1) - HR_i^-$, and the post - pre difference in that same time period had the region not received the military intervention, $Y_i(0) = HR_i^+(0) - HR_i^-$, (see Figure 2 for a sketch of the definition of the modified potential outcomes). Nevertheless, the estimand defined with Y is identical to the estimand previously defined with HR because $Y_i(1) - Y_i(0) = HR_i^+(1) - HR_i^+(0)$. Thus,

$$\tau = \frac{\sum_{i=1}^{N} HR_i^+(1) - HR_i^+(0)}{N} = \frac{\sum_{i=1}^{N} Y_i(1) - Y_i(0)}{N} = \overline{Y}(1) - \overline{Y}(0).$$

Therefore, we wish to estimate the post - pre homicide rate difference in the treated regions had they not received a military intervention (i.e., the missing potential outcome $Y_i(0)$). These missing potential outcomes are estimated using municipalities that were not intervened but are similar to those forming the treated regions before 2007, in particular in their 2006 homicide rates, but might not be contiguous. We believe this makes sense because it is reasonable to assume that in the absence of an intervention there are no spill

over effects on the neighbors. Furthermore, we believe that constructing hypothetical regions similar to the actual treated ones is better achieved by the matching method that we used than by requiring geographical contiguity because many key covariates relevant for the outcomes, like pre-intervention homicide rates, political composition and measures of pre-intervention government presence (for example percentage of indigenous language speakers, average education level and road network longitude), are unlikely to be well matched if a contiguity restriction is imposed. We believe that the importance of these variables trumps the remaining relationships between contiguous municipalities that are not controlled using the covariates considered. We denote the *observed* outcome of treated region i by $Y_i(1)$, and its *estimated* outcome under control by $\tilde{Y}_i(0)$.

We estimate the intervention effects by *multiply imputing* the missing potential outcome for each treated region. We create a donor pool of hypothetical control regions to impute these outcomes. The number of municipalities forming each treated region, N_i, determines the size of the corresponding donor pool, which is 2^{N_i}. The donor pool sizes range from 2^4, for the Tijuana region, to 2^{36}, for the Acapulco region. Instead of listing all hypothetical control regions and then drawing from this pool at random, for every imputation of $Y_i(0)$, we generate a hypothetical control region by assuming comparable quality between matches for any municipality, and randomly selecting a match for every municipality in treated region i. Let $k_j \sim 1 + Bern(1/2)$ denote the index of the match selected for the jth municipality in treated region i. Analogous to the notation defined in Subsection 3.3 and the beginning of this one, let $H_{ijk_j}^-$, $P_{ijk_j}^-$, $H_{ijk_j}^+$ and $P_{ijk_j}^+$ denote the homicide count and the population of this match one year before (-) and after (+) the first intervention in the region. Thus one imputation (draw) of $Y_i(0)$ is obtained in the following way,

$$\widetilde{Y}_i^{draw}(0) = \frac{\sum_{j=1}^{N_i} H_{ijk_j}^+}{\sum_{j=1}^{N_i} P_{ijk_j}^+} - \frac{\sum_{j=1}^{N_i} H_{ijk_j}^-}{\sum_{j=1}^{N_i} P_{ijk_j}^-}.$$

17

We use $\widetilde{Y}_i^{draw}(0)$ to represent a drawn value (imputation) of the intervention effect for region i,

$$\tau_i^{draw} = Y_i(1) - \widetilde{Y}_i^{draw}(0).$$

Then, we obtain a draw of the estimand by averaging the draws for the N treated regions,

$$\tau^{draw} = \frac{\sum_i \tau_i^{draw}}{N}.$$

Repeating this imputation multiple times gives us an approximation to the posterior distribution of the estimands under our assumptions, and a measure of uncertainty for the point estimates of each region, as well as for the overall average. Let $\hat{\tau}$ and $\hat{\tau}_i$ be the mean of the distributions of τ^{draw} and τ_i^{draw}, i.e., our point estimates. The 95% intervals shown in Figure 3 and Table 4 are calculated by finding the 0.025 and 0.975 quantiles of the distributions of these draws. Note that the RCM allows for the region-level effect assessment, although these estimations are more imprecise than that of the average treatment effect on the treated regions.

Recall that the use of Y_i instead of HR_i only has an effect in the estimation and not on the estimand. Therefore, the use of Y_i terminology (*pre - post difference*) is restricted to this section. Henceforth, all results are given in terms of the post intervention homicide rates.

5   Results

Understanding effectiveness as a decrease in *violence*, measured by homicide rate, our analysis indicates that the military interventions were *ineffective*, resulting in an increase in the average homicide rate. However, the estimated effect varies considerably across the treated regions. Our results are displayed in Figures 3 and 4, and Table 4. Out of the eighteen treated regions, only Rincón de Romos has an estimated drop in the homicide rate post intervention,

relative to what it would have been without the intervention, where zero is not included in the 95% imputation interval of the average effect. Nevertheless, for the Apatzingán and La Piedad regions, the intervention is estimated to have decreased the homicide rate, but the resulting 95% intervals do contain zero. The Juárez region is a clear outlier, and its estimate of the intervention effect is almost three times as large as the one closest to it in terms of the estimated intervention effect. Removing Juárez from the analysis yields a smaller estimate of the average treatment effect on the treated, but the interval still excludes 0.
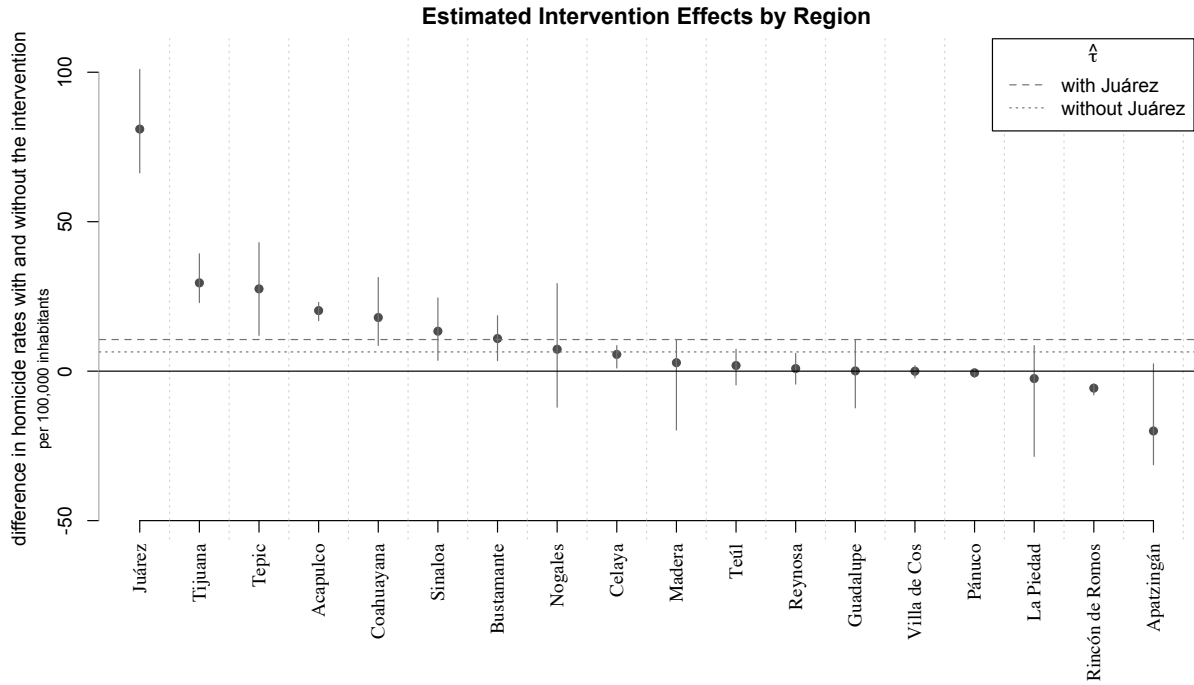

Figure 3: Results, shown in decreasing order, of the estimated treatment effect per region τ_i - the comparison between the observed one year post intervention homicide rate and the estimate of the homicide rate that same year without the intervention. The dark horizontal line denotes zero, and the dashed and dotted lines denote respectively the average intervention effect including and excluding Juárez.

Assessments for Tijuana, Nogales and Apatzingán regions should be made with care given the limitations in the overlap shown in the propensity score distributions of municipalities in each of these treated regions and their matched controls. However, it is unlikely that the direction of the estimated effect for the Apatzingán region would change given how extreme
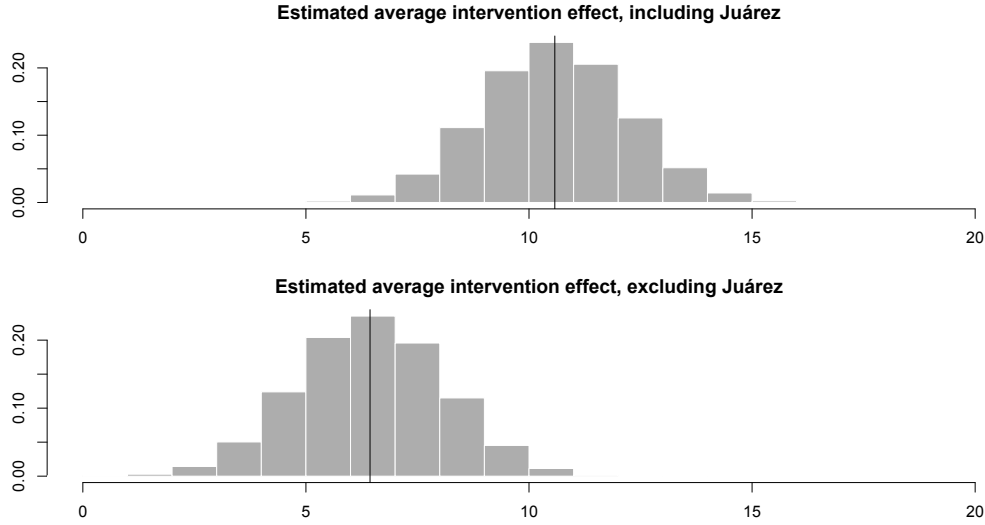
Figure 4: Aggregate results showing distributions of the estimated average difference in post - pre homicide rates with and without the intervention, including (top) and excluding (bottom) Juárez.

the observed drop in homicide rate for this region is (see Figure 5); only three observed 2008 - 2006 homicide rate differences were lower.

An obvious limitation of our analysis is that the cartel information was not used. Even though matching improved balance on criminal "rivalry" related homicide rates, the difference between municipalities in treated and hypothetical control regions is still considerable, and the difference in means for 2006 State GDP increased after matching. If this variable is crucial for unconfoundedness, a causal conclusion is not valid but a conditional association interpretation of these results can still be made.

Furthermore, we do not have a comprehensive list of interventions. Such a list can only remove municipalities available for matching from the current control pool. Because the interventions took place in municipalities that were thought to need them the most, we would expect lower pre-intervention homicide rates in the *true-controls*, leading to a larger difference between control and treated pre-intervention homicide rates compared to that obtained now, which might include some *intervened but identified-as-control* municipalities

20

Extremeness of drop in homicide rate in the Apatzingán region

Observed differences in 2008 - 2006 homicide rates for control municipalities
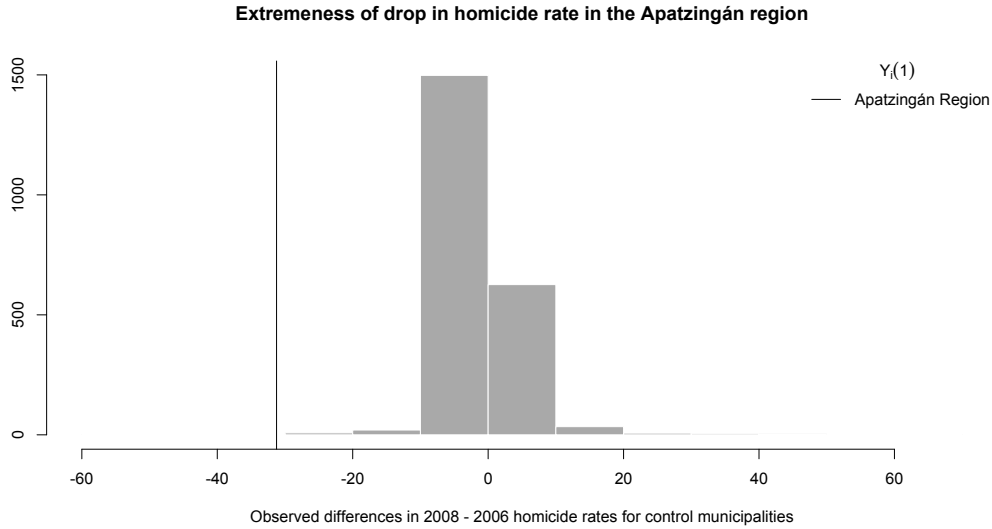
Figure 5: Comparison of $Y_i(1)$ for the Apatzingán region relative to all the observed 2008 - 2006 homicide rate differences for all the municipalities in the control pool.

in the control pool. Thus, assuming pre-intervention homicide rates are predictive of the post-intervention homicide rates, using these *true-control* matches would most likely increase the estimated difference between the homicide rate with the intervention and the homicide rate without the intervention, relative to the one reported in this paper.

Perhaps confining the analysis to one year post-intervention is too limited to evaluate the effectiveness of the military interventions fairly. The homicide rate in Ciudad Juárez, located in the region with the highest increase in homicide rate, dropped considerably in 2012. [8] This suggests that following the homicide rate trend several years after the intervention takes place will give a more complete picture of the effect of an intervention. It would not be surprising that initially a military intervention overwhelms the local balances in power, cartels and local police - which is particularly relevant if these bodies were infiltrated by organized crime, increasing the violence in the short term; but that military presence is proven to be effective in a longer period of time. We do not report these estimates beyond

---

[8]http://mexico.cnn.com/nacional/2013/01/05/los-homicidios-disminuyen-mas-del-50-en-ciudad-juarez-de-2011-a-2012

Region	Number of Municipalities	Date of First Intervention	Effect One Year After		
			$Y_i(1)$	$\hat{\tau}_i$	95% interval
Juárez	16	2008	117.96	80.97	(66.34, 100.95)
Tijuana	4	2008	45.20	29.53	(22.93, 39.25)
Tepic	11	2010	42.98	27.57	(11.96, 43.02)
Acapulco	36	2008	26.57	20.26	(16.89, 22.97)
Coahuayana	6	2010	23.03	17.94	(8.60, 31.33)
Sinaloa	28	2007	11.58	13.36	(3.61, 24.49)
Bustamante	5	2010	14.42	10.89	(3.49, 18.54)
Nogales	6	2008	40.15	7.31	(-12.06, 29.30)
Celaya	9	2009	5.24	5.62	(1.06, 8.57)
Madera	12	2010	3.09	2.82	(-19.70, 10.51)
Teúl	10	2009	5.62	1.90	(-4.60, 7.35)
Reynosa	24	2008	4.86	0.87	(-4.35, 5.94)
Guadalupe	20	2009	13.11	0.08	(-12.28, 10.43)
Villa de Cos	22	2008	4.54	-0.02	(-2.25, 1.82)
Pánuco	14	2007	-0.02	-0.55	(-2.13, 0.48)
La Piedad	9	2010	9.94	-2.45	(-28.51, 8.56)
Rincón de Romos	7	2008	-0.82	-5.68	(-7.95, -4.30)
Apatzingán	9	2007	-31.25	-20.05	(-31.32, 2.44)
Average (including Juárez)	248	-	18.68	10.58	(7.42, 13.72)
Average (excluding Juárez)	232	-	12.84	6.44	(3.25, 9.59)

Table 4: Point estimates and 95% intervals.

the first year because we only have homicide counts up to 2011, reducing considerably the number of regions that can be used to estimate the second year post intervention effect and beyond.

6   Conclusion

The design and analysis stages of this observational study posed interesting challenges. Here we explored the simplest approach that we believed was possibly appropriate. However, this type of study is probably better suited for a causal network analysis because of the interference between municipalities and the possibility of defining multiple levels of treatment. Further exploration of such an approach is appealing for future work because the causal

network area of statistics is underdeveloped and current approaches are unsatisfactory.

In the no-interference approach we took here, the feasibility of SUTVA induced a modification of the definition of the units. Furthermore, the imputation procedure involved subtleties in the construction of hypothetical control regions "comparable" to the treated ones. We proposed an analysis method based on permutation ideas that allowed for the multiple imputation of the homicide rate of the treated regions, had they not received a military intervention, by creating donor pools of comparable hypothetical control regions. Our approach is more principled than those offered by Escalante (2011) and Merino (2011), increasing the feasibility of the implicit assumptions being made and the explicit assessment of covariate balance. Our method multiply imputes the missing potential outcomes at the region level resulting in approximate distributions for the treatment effects at the regional level and for the average treatment effect on the treated regions. This enables the comparison of effects between different treated regions.

There are limitations (previously discussed) in our study that might affect the feasibility of the unconfoundedness assumption. Nevertheless, if unconfoundedness is not achieved, unlike for the previous studies, a principled conditional association interpretation of these results can still be made and may be interesting.

References

Ahmed, A., Husain, A., Love, T. E., Gambassi, G., Dell'Italia, L. J., Francis, G. S., Gheorghiade, M., Allman, R. M., Meleth, S., and Bourge, R. C. (2006), "Heart failure, chronic diuretic use, and increase in mortality and hospitalization: an observational study using propensity score methods," *European heart journal*, 27, 1431–1439.

Coscia, M. and Rios, V. (2012), "Knowing where and how criminal organizations operate using web content," .

Escalante, F. (2011), "Homicidios 2008-2009: La muerte tiene permiso," *Revista Nexos.*

Merino, J. (2011), "Los operativos conjuntos y la tasa de homicidios: una medición," *Revista Nexos.*

Rosenbaum, P. R. and Rubin, D. B. (1983), "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41–55.

Rubin, D. B. (1974), "Estimating causal effects of treatments in randomized and nonrandomized studies," *Journal of educational Psychology*, 66, 688.

— (2007), "The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials," *Statistics in medicine*, 26, 20–36.

— (2008), "For objective causal inference, design trumps analysis," *The Annals of Applied Statistics*, 808–840.