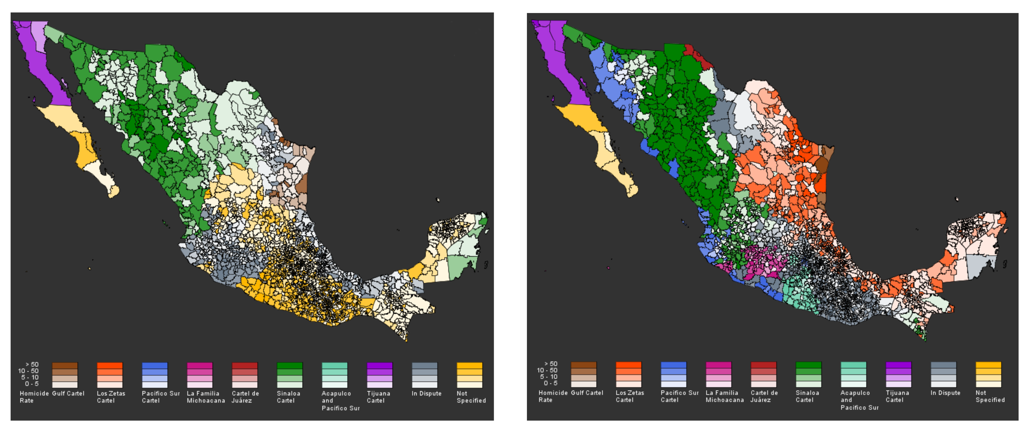
Visualization and Causal Inference of the Mexican Drug War

Valeria Espinosa and Joseph Kelly

vespinos@fas.harvard.edu, kelly2@fas.harvard.edu

Statistics Department, Harvard University

Visualize the problem



(a) 2007 (b) 2010

We attempt to answer whether **homicide rates increase significantly after a military intervention**.

Estimand

Let $Y_i(1)$ and $Y_i(0)$ denote the change in homicide rate of region j one year after receiving a military intervention, and what it would have been if it hadn't received it (Rubin Causal Model). Our estimand is the average causal effect of the military intervention for the regions that were intervened. That is

$$\tau = Y(1) - Y(0) = \frac{\Sigma_{i=1}^{I} Y_i(1) - Y_i(0)}{I}.$$

Let N_i denote the number of municipalities that correspond to region i, then

$$Y_i(1) = \sum_{i=1}^{N_j} w_{ij} Y_{ij}(1) \text{ and } Y_i(0) = \sum_{j=1}^{N_i} w_{ij} Y_{ij}(0),$$

$$\text{where } w_{ij} = \frac{\text{Pop}_{ij}}{\text{Pop}_i} \text{ and } \text{Pop}_i = \sum_{j}^{N_i} \text{Pop}_{ij}.$$

However, $Y_i(0)$ and $Y_{ij}(0)$ are missing $\forall i, j$.

Estimation

Propensity score matching was used to identify a set of 5 control municipalities that look like each treated ones, and ultimately estimate $Y_{ij}(0)$ and $Y_i(0)$. Let M_{ij} be the number of municipalities matched to the jth municipality in region i, and

$$\text{PopM}_{ij} = \sum_{k=1}^{M_{ij}} \text{PopM}_{ijk}$$

is the sum of their populations. Then,

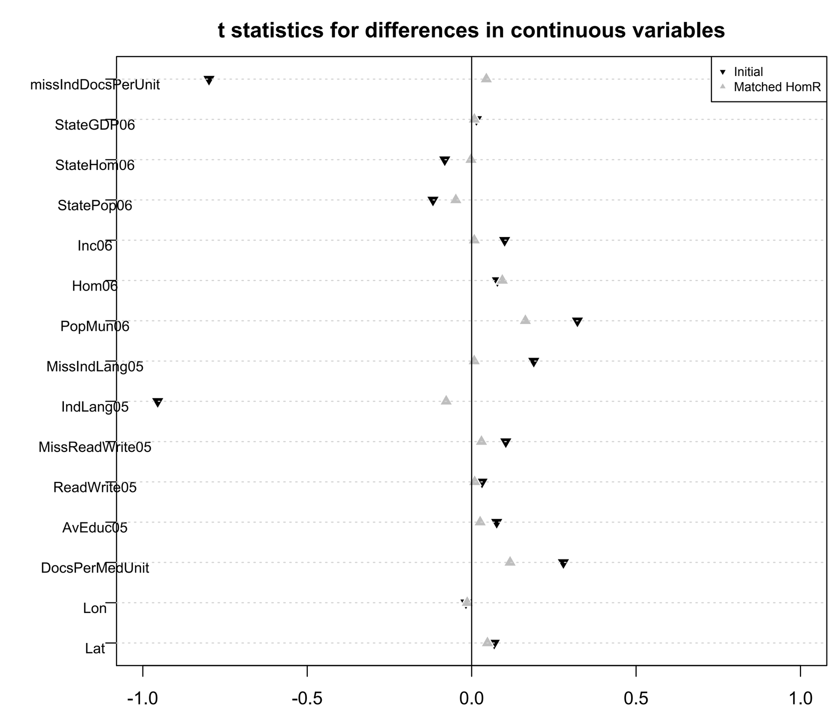
$$\hat{Y}_{ij}(0) = \sum_{k=1}^{M_{ij}} v_{ijk} Y_{ijk}(0),$$

where $v_{ijk} = \frac{\text{PopM}_{ijk}}{\text{PopM}_{ij}}$. Therefore,

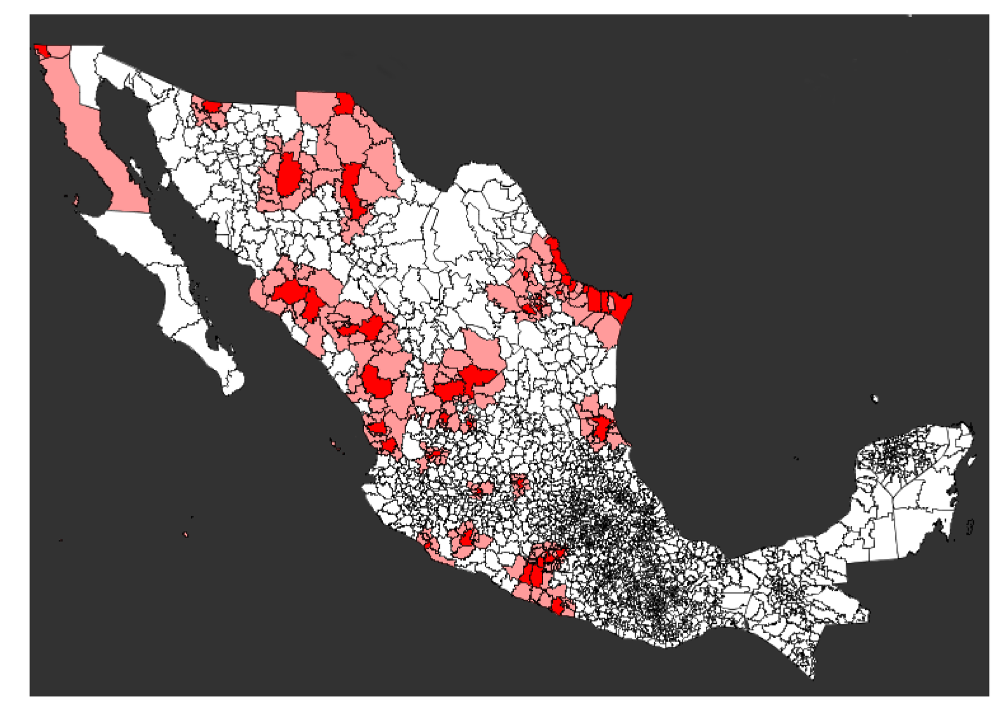
$$\hat{Y}_i(0) = \sum_{j=1}^{N_i} w_{ij} \hat{Y}_{ij}(0) = \sum_{j=1}^{N_i} w_{ij} \sum_{k=1}^{M_{ij}} v_{ijk} Y_{ijk}(0),$$

and

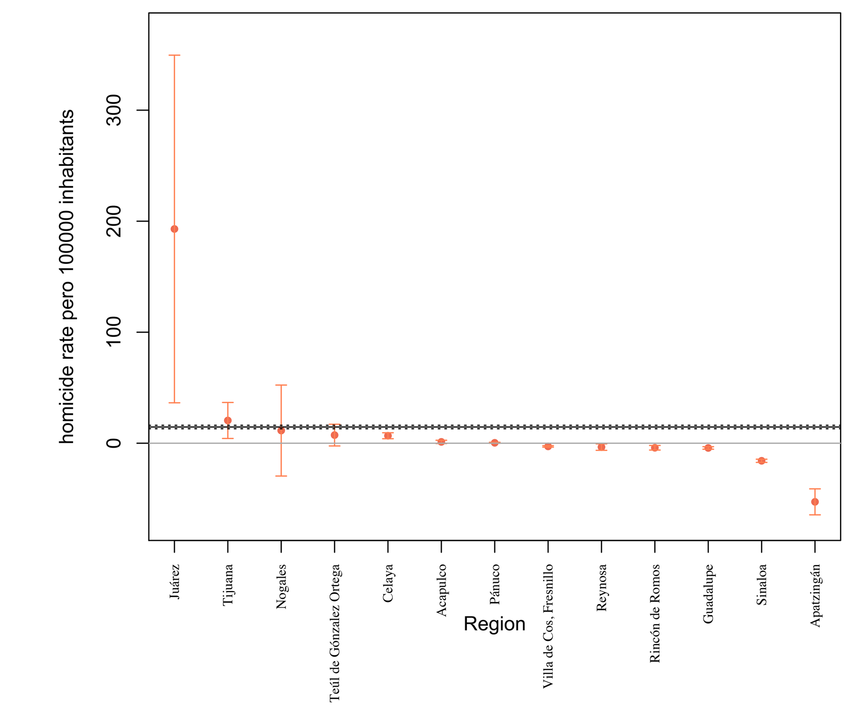
Visualization



(c) Love plot - balance checks



(d) Interventions and SUTVA



(e) Results

- **SUTVA** - stable unit treatment value assumption
  - **No hidden values of treatments** A broad definition of what "military intervention"- any mili means in this context helps us think of a two level treatment: receiving a military intervention (defined as ... see paper(that have resulted in deaths?)) or not receiving it.
  - **No interference between units** The main idea is to group neighboring regions that have received military interventions in such a way that distances make the "no interference" assumption more reasonable. For treated regions that are side to side were also assessed in terms of neighboring geographic situation such as lack of highways connecting them The last homicide rate that we have corresponds to 2010. That eliminated some of the interventions mentioned in the Nexos paper. Following this reasoning 2213 municipalities were included in the initial control pool. Our 13 treated regions are:

    update this table to have the number of municipalities that received the intervention AND the main names. Put the image of the intervention map next to it
- **Unconfoundedness**: Unfortunately we didn't get experts to guide most of our decisions. However, we did get to interact with a couple of them and made our covariate choices based on the information received and our understanding of the relevant information. Our covariates include: location, political party before Calderón, income 2006, education and medical information at 2005, percentage of indigenous language speakers, 2006 homicide rate at the municipality level, and GDP, Homicide Rate and Population at the state level.
- **Missing Data:** there were few treated units that had one covariate (Doctors per medical unit) missing, we exactly matched on that and Political Party in power before Calderón.

Results

unit	Region	number of municipalities	Date of first intervention	Within Region Effect	SD Neyman
1	**Tijuana**	5	2008	20.49	8.27
2	**Nogales**	5	2008	11.41	20.90
4	**Juárez**	15	2009	192.99	79.88
5	**Pánuco**	14	2007	0.37	0.24
6	**Reynosa**	24	2008	-3.49	1.48
8	**Guadalupe**	20	2009	-4.27	0.58
9	**Villa de Cos, Fresnillo**	18	2008	-2.87	0.34
10	**Teúl de González Ortega**	10	2009	7.32	4.99
11	**Rincón de Romos**	8	2008	-4.10	1.05
12	**Sinaloa, Badiraguato, Pueblo Nuevo**	27	2007	-15.84	0.74
15	**Celaya**	9	2009	6.74	1.37
16	**Apatzingán**	10	2007	-52.81	5.97
18	**Acapulco**	35	2008	1.19	0.77
	Average	205	-	14.61	0.79

Key References & Data Source

[1] Abadie Synthetic Matching

[2] Imbens G. & Rubin D.R., (2012)

[3] Rubin D.R. ,*Matched Sampling for Causal Effects*,

[4] Diego Valle visualization

[5] CIDAC

[6] INEGI

[7] Stratfor Maps