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

Abstract

In Mexico, the presidency of Felipe Calderón (2006-2012) has been characterized for the war against organized crime, raising many questions regarding security and violence. The main question of interest is: **do homicide rates increase significantly after a military intervention?**. Due to the observational nature we explored the feasibility of causal inference for the data obtained. There are many challenges involved in answering this question. Here, we attempt to point them out and a first try at answering this question. As any good observational study, which mimics a randomized experiment, the *design* and *analysis* steps are clearly separated. The de

Estimand

Following the Rubin Causal Model, let $Y_j(1)$ and $Y_j(0)$ denote the homicide rate of region j one year after it received a military intervention, and what it would have been at that same point in time if it hadn't received the military intervention. The estimand of interest is the average causal effect of the military intervention for the regions that were intervened. That is

$$\tau = Y(1) - Y(0) = \frac{\sum_j Y_j(1) - Y_j(0)}{J}.$$

For this case, we can think of the treated regions as the population of interest because in principle only regions like those are likely to receive a military intervention. We assume $Y_j(1)$ is observed for all treated units. Let N_j denote the number of municipalities that correspond to region j, then

$$Y_j(1) = \sum_{i=1}^{N_j} w_{ij} Y_{ij}(1),$$

where $\text{Pop}_j = \sum_i^{N_j} \text{Pop}_{ij}$, and

$$w_{ij} = \frac{\text{Pop}_{ij}}{\text{Pop}_j}.$$

However, $Y_j(0)$ is missing for all j. Following the reasoning above,

$$Y_j(0) = \sum_{i=1}^{N_j} w_{ij} Y_{ij}(0),$$

and all $Y_{ij}(0)$ are unobserved.

Assumptions & Challenges

- **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:

unit	number of municipalities	Date of first intervention	**unit**	number of municipalities	Date of first intervention
1	5	2008	**10**	10	2009
2	5	2008	**11**	8	2008
4	15	2009	**12**	27	2007
5	14	2007	**15**	9	2009
6	24	2008	**16**	10	2007
8	20	2009	**18**	35	2008
9	18	2008			

  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

Estimation

We attempt to use the information of all other municipalities to estimate each $Y_{ij}(0)$ to obtain an estimate $Y_j(0)$. We use propensity score matching to identify a set of control municipalities that look like the treated ones.

To follow the guidelines for observational studies we will first clarify what the analysis protocol will be. Given that weights will be given, this specification will determine the way the love plots are created during the balance check process. Let M_{ij} be the number of municipalities matched to the ith municipality in region j. Let

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

denote the total population of all M_{ij} municipalities matched to the ith treated municipality in region j. 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,

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

and

Others

The "alert" block environment looks like this. It also has justified text, but it has a border and a light background to make it stand out. You can create one like so:

```
\begin{alertblock}{Title}
.......
\end{alertblock}
```

Alert Block Colours

You can similarly modify the colours for alert blocks (but try not to overdo it):

```
\setbeamercolor{block title}
{fg=black,bg=norange}
\setbeamercolor{block body}
{fg=black,bg=white}
```

Key References

[1] Abadie Synthetic Matching

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

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

[4] Nexos paper