

The Effect of Conditional Cash Transfer Policies on Regional Crime Levels:

Evidence from a Synthetic Controls Framework

Felipe Santos-Marquez

<https://felipe-santos.rbind.io>

Research Assistant / PhD Student
Chair of International Economics
Technische Universität Dresden
Germany

Prepared for the Brown Bag Seminar July 7th 2021

[slides available at: <https://bbs-felipe.netlify.app>]

Motivation:

- large regional inequality between Colombian municipalities and high homicide rates.
- There is no certainty over the effect of conditional transfers on violent crime, and especially on homicides.
- Scarce academic literature on the impact of CCT programs on crime at the municipal level in Colombia.

Research Objective:

- Which are determinants of homicide rates for Colombian municipalities?
- **To what extent the coverage of conditional cash transfer program in Colombia (the pacific region of Colombia) may affect homicide rates?.**

Methods:

- Bayesian Model Averaging **BMA** (Fernandez et al. (2001)).
- Synthetic control methods (Abadie and Gardeazabal (2003)).

Data:

- Municipal panel dataset CEDE, released by the University of The Andes.
- National Administrative Department of Statistics.

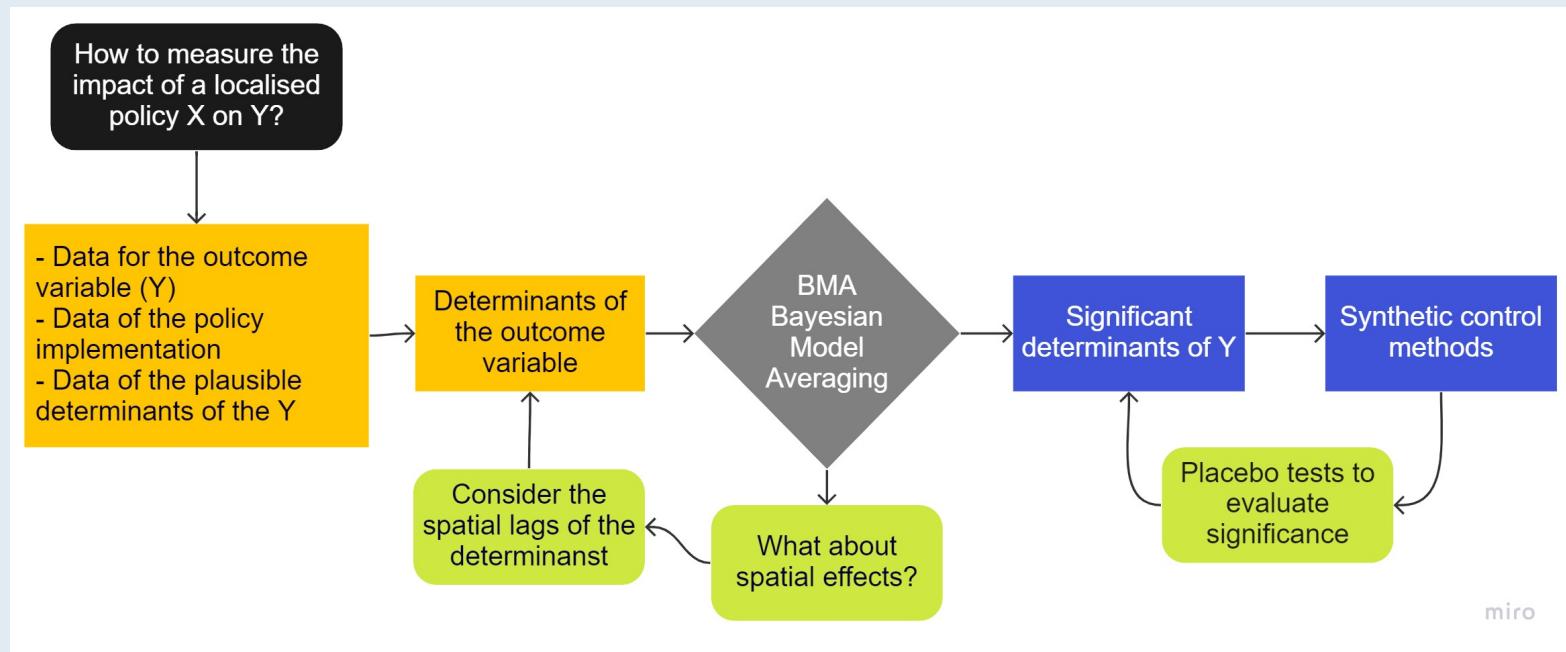
Main Results:

- 15 variables are found to be important determinants of homicide levels. They are related to **crime, inequality, drug-trafficking, conflict and literacy**.
- The importance of spatial effects is highlighted by the fact that out of 15 variables **9 are spatially lagged variables**.
- It was reported that by 2018, **the average homicide rates were lower for high CCTs coverage municipalities** when compared to synthetic controls ("copies" made out of a pool of low CCT coverage municipalities).

Outline of this presentation

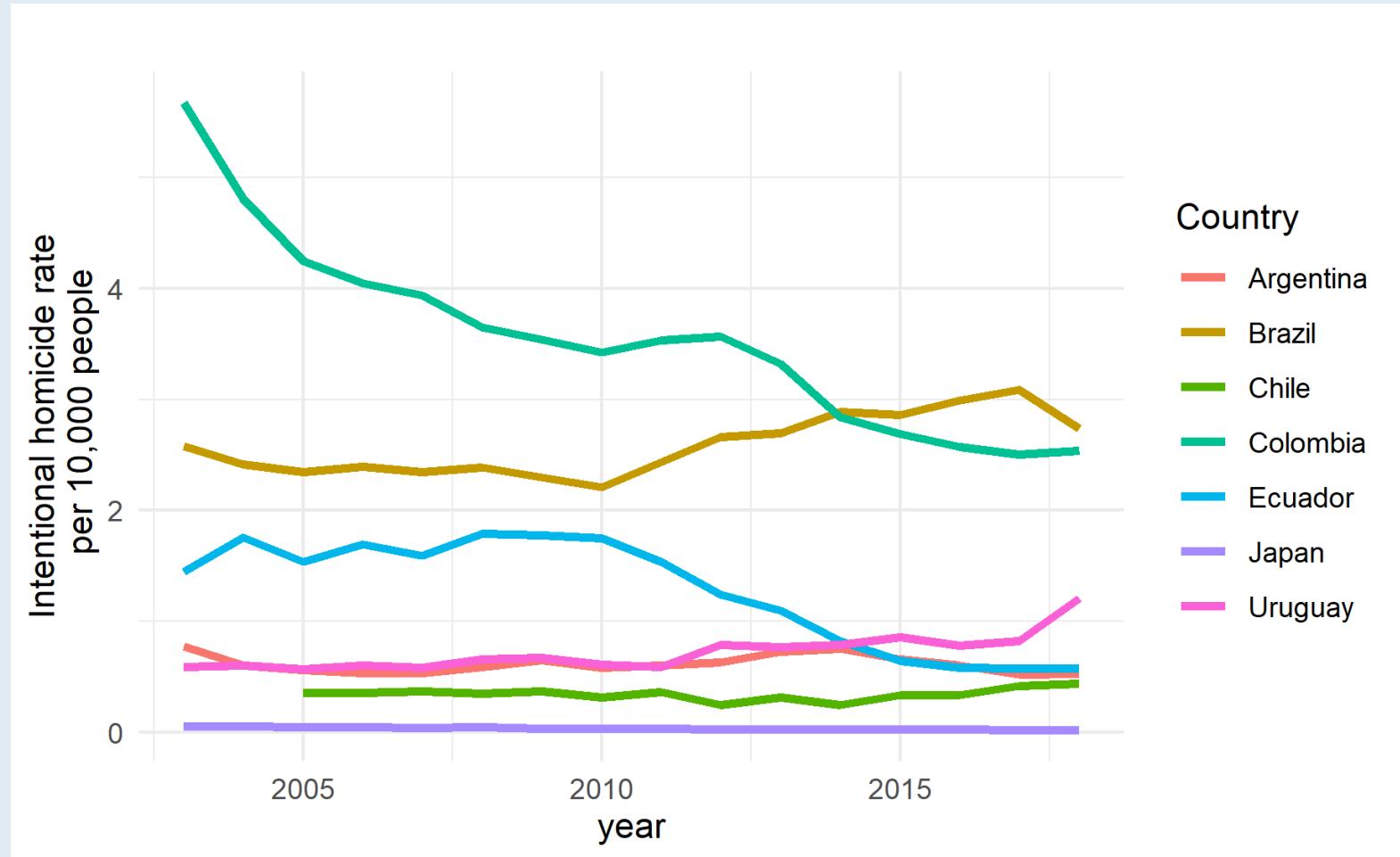
1. A visual representation of the project
2. Regional disparities and crime in Colombia
3. What are CCTs?
4. BMA Bayesian Model aveaging
 - Methodology
 - Results
5. Synthetic Controls Methods
 - Methodology
 - Results
6. Conclusions
7. Open Science

1. The paper in one chart



$BMA \rightarrow \text{determinants of } Y \rightarrow \text{Synthetic Controls} \rightarrow \text{The impact of } X \text{ on } Y$

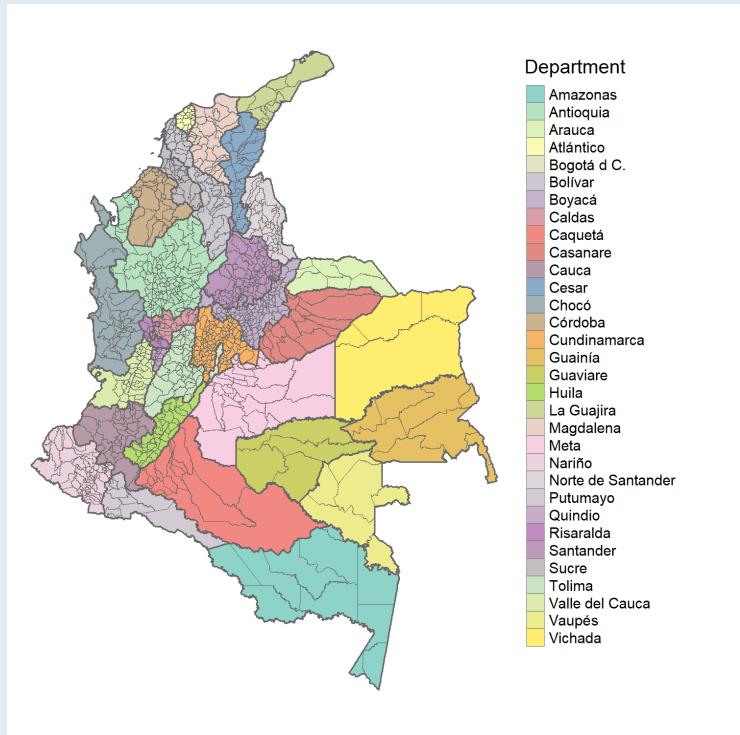
2. Homicide rates over time



Intentional homicides for selected South American countries and Japan (Source: Author's calculations using data from the WDI World Bank (2020))

Colombian administrative levels

States and Municipalities

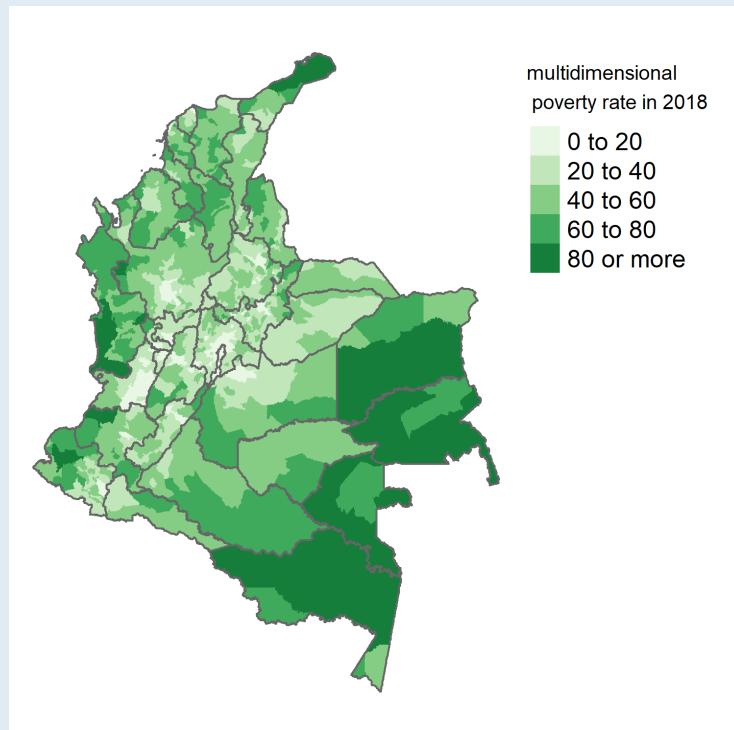


Natural Regions

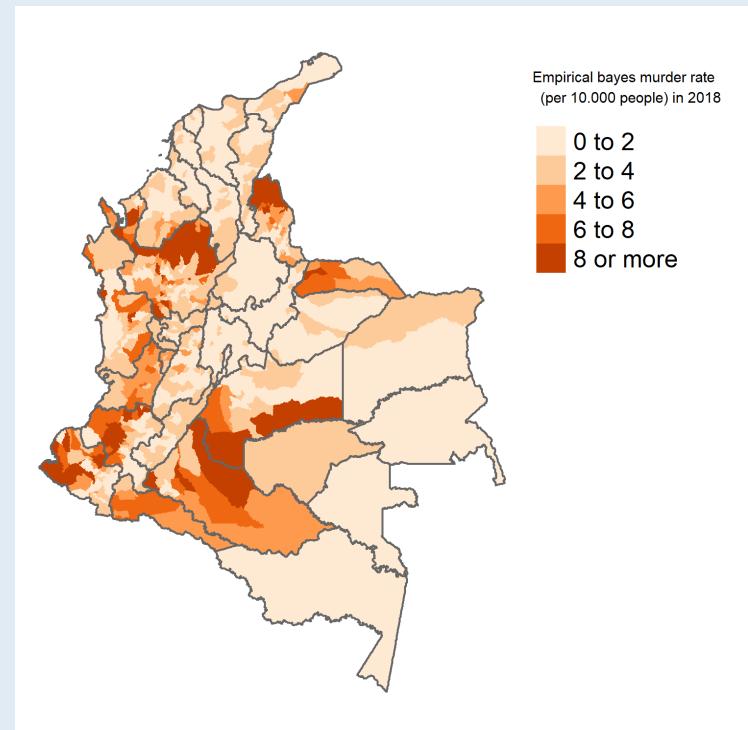


2. Large regional disparities in Colombia

Well-being



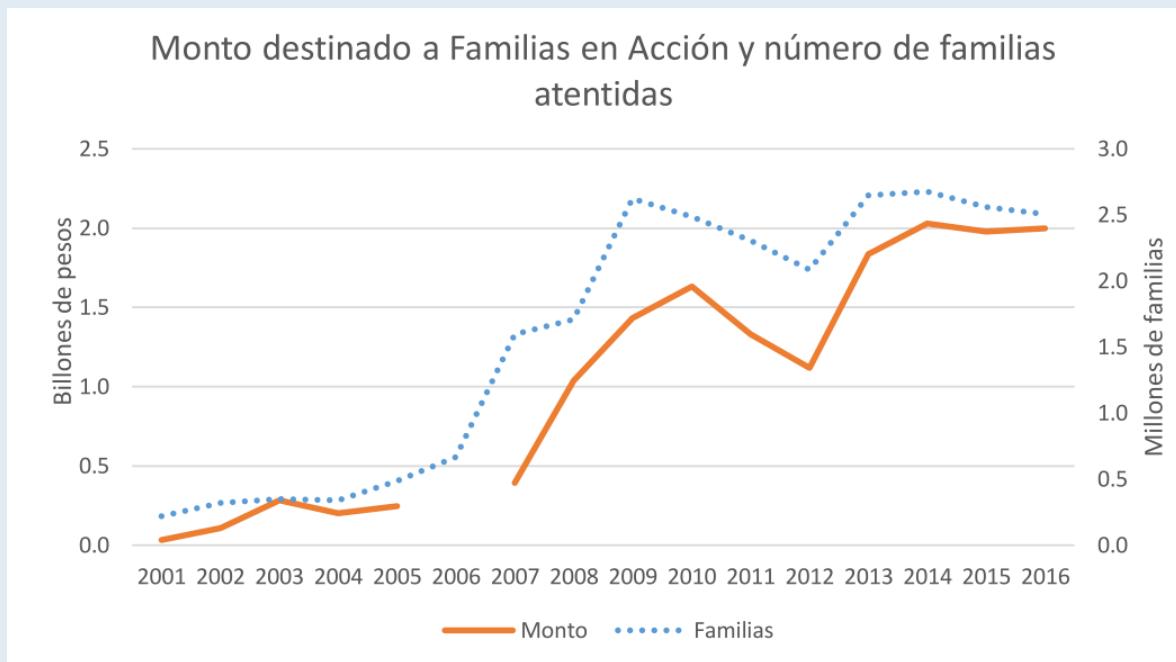
Crime



(In Germany = about 0.1 per 10.000
people)

3. Conditional Cash Transfers

- In general terms, the CCTs programs consisted on monetary transfers in order to encourage the poorest Colombian families to have children and teenagers attend schools, making sure their diet and health wouldn't worsen during the economic recession. ([Urrutia and Robles 2018](#))
- It has become a permanent public policy, several studies have been focused on the implications of this policy. ([Urrutia and Robles 2018](#))



4. BMS ---- In the economics literature (to start with)

I Just Ran Two Million Regressions (Sala-I-Martin 1997)

$$(2) \quad \gamma = \alpha_j + \beta_{yj}y + \beta_{zj}z + \beta_{xj}x_j + \varepsilon$$

where y is a vector of variables that always appear in the regressions (in the Levine and Renelt paper, these variables are the initial level of income, the investment rate, the secondary school enrollment rate, and the rate of population growth), z is the variable of interest, and $x_j \in X$ is a vector of up to three variables taken from the pool X of N variables available. One needs to estimate this

problems, I decided to follow Levine and Renelt and allow all the models to include three fixed variables, so when I combine these three variables along with the tested variable and then with trios of the remaining 59 variables, I always have regressions with seven explanatory variables.

Of all the variables in the literature. I chose a total of 62. The selection was made keeping in mind that I want variables that are measured at the beginning of the period (which is 1960) or as close as possible to it to minimize endogeneity. This eliminated all those variables that were computed for the later years only.

TABLE I—MAIN RESULTS OF REGRESSIONS
(DEPENDENT VARIABLE = GROWTH)

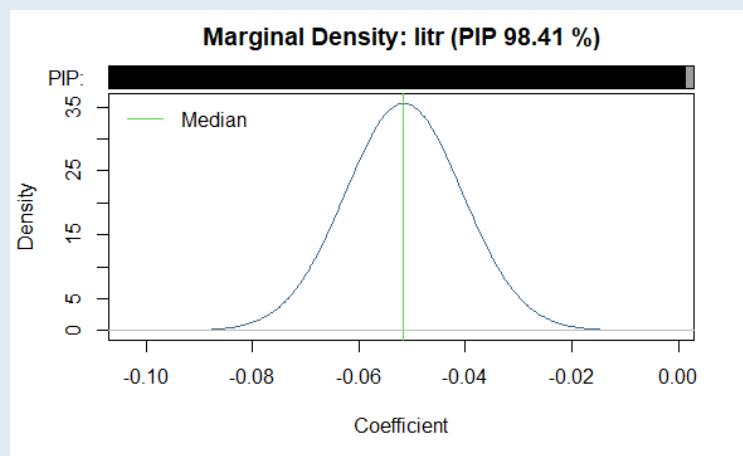
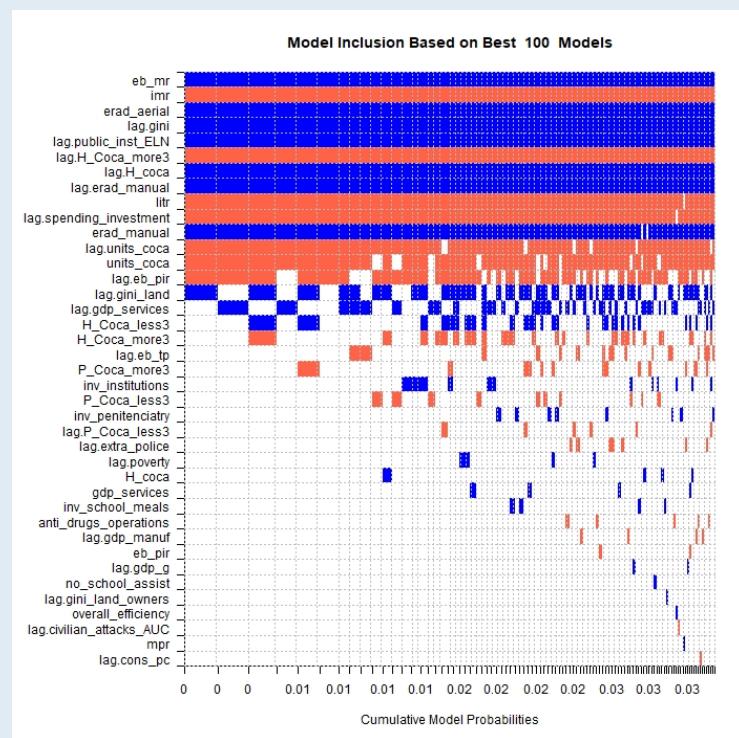
Independent variable	(i) β	(ii) SD	(iii) CDF ^a
Equipment investment	0.2175	0.0408	1.000
Number of years open economy	0.0195	0.0042	1.000
Fraction Confucian	0.0676	0.0149	1.000
Rule of law	0.0190	0.0049	1.000
Fraction Muslim	0.0142	0.0035	1.000
Political rights	-0.0026	0.0009	0.998
Latin America dummy	-0.0115	0.0029	0.998
Sub-Saharan Africa dummy	-0.0121	0.0032	0.997
Civil liberties	-0.0029	0.0010	0.997
Revolutions and coups	-0.0118	0.0045	0.995
Fraction of GDP in mining	0.0353	0.0138	0.994
SD black-market premium	-0.0290	0.0118	0.993
Primary exports in 1970	-0.0140	0.0053	0.990
Degree of capitalism	0.0018	0.0008	0.987
War dummy	-0.0056	0.0023	0.984
Non-equipment investment	0.0562	0.0242	0.982
Absolute latitude	0.0002	0.0001	0.980
Exchange-rate distortions	-0.0590	0.0302	0.968
Fraction Protestant	-0.0129	0.0053	0.966
Fraction Buddhist	0.0148	0.0076	0.964
Fraction Catholic	-0.0089	0.0034	0.963
Spanish colony	-0.0065	0.0032	0.938

^a Nonnormal.

4. Bayesian Model Averaging BMA - methods and results:

$$y = \alpha_i + X_i \beta_i + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I)$$

- **how can researchers select just a handful of determinants?**
- **how to evaluate the importance of the inclusion of specific determinants in the model?**
- Bayesian Model Averaging (BMA) methods attempt to overcome these problems by estimating linear models for all (**MANY**) possible combinations of determinants X_i



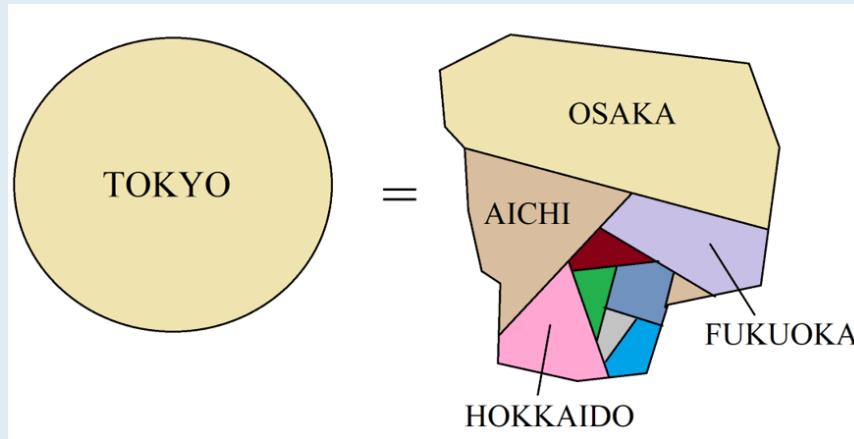
4. 212 variables between original and spatially lagged variables were tested as determinants of 2018 homicide rates.

After running 2 million regressions...

description	PIP all variables	Post mean all
murder rate	1.000	0.356 (0.031)
infant mortality rate	1.000	-0.058 (0.011)
Coca H aerial eradication	1.000	0.001 (0)
Lag. Coca crops with more than 3 H	1.000	-0.041 (0.006)
Lag. Coca hectares	1.000	0.023 (0.003)
Lag Attacks against public institutions by rebels	1.000	36.062 (7.46)
Lag. H manual eradication	1.000	0.004 (0.001)
Lag. Income Gini	1.000	0.25 (0.041)
H manual eradication	0.990	0.002 (0.001)
Lag. Spending on investment local government	0.986	-0.095 (0.025)
Literacy rate	0.984	-0.051 (0.013)
Number of land units with coca crops	0.961	-0.004 (0.003)
Lag. Number of land units with coca crops	0.943	-0.015 (0.004)
Lag. Personal injury rates	0.681	-0.034 (0.025)
Lag. Land Gini	0.564	0.022 (0.021)

5. Synthetic control methods

visual intuition (In terms of GDP per capita)



$$TOKYO = 0.4 * OSAKA + 0.2 * AICHI + 0.1 * FUKUOKA + \dots$$

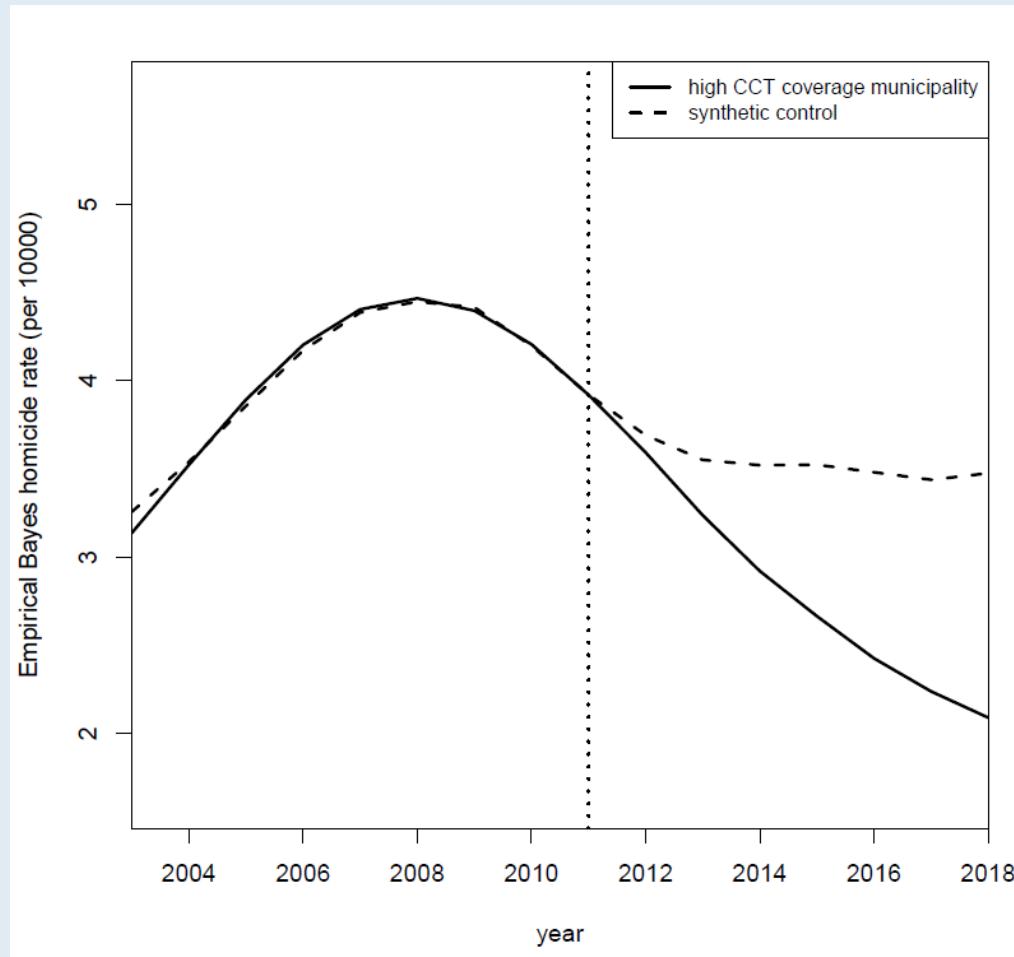
In terms of crime

The weights are found so that the synthetic municipality has a similar crime trend compared to the treatment municipality (2003-2011) and similar determinants of crime.

5. Results: Synthetic control methods

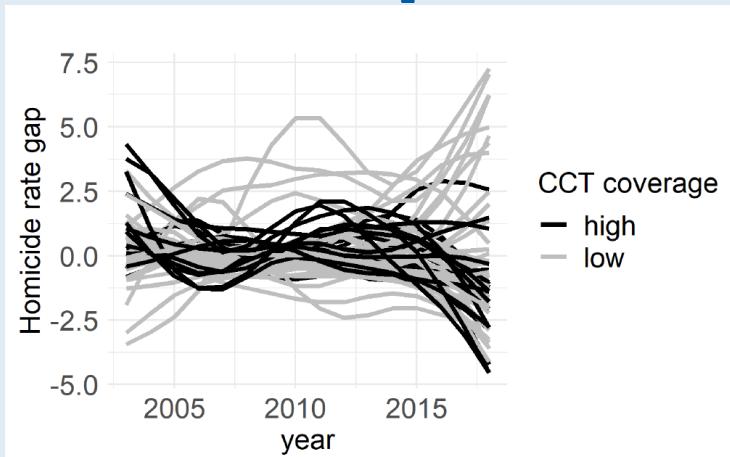
high – CCT – coverage > 70%

low – CCT – coverage < 30%

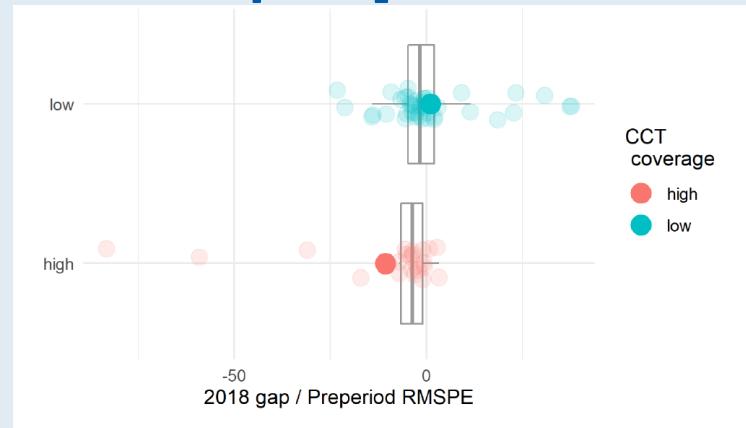


5. The effect of CCTs on crime

Crime gaps for treatment municipalities
and control placebos



Overall effects = the gap in 2018 / Root mean squared predicted error



$$\text{RMSPE} = \left(\frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right)^2 \right)^{1/2}$$

A t-test shows that the mean effect (lower crime) is statistically lower for the treatment group.

5. Robustness Checks

Pop	high CCT%	#regions	low CCT%	#regions	Postperiod RMSPE / Preperiod RMSPE	2018 gap	2018 gap / Preperiod RMSPE
larger than 10K	65	31	35	48	0.7	0.8	0.1
	70	22	30	41	0.5	0.04	0.03
	75	18	25	36	0.05	0.1	0.4
	80	14	20	28	0.5	0.3	0.2
	85	13	15	24	0.4	0.05	0.3
	90	12	10	18	0.4	0.4	0.3
all	65	43	35	70	0.4	0.8	0.09
	70	32	30	60	0.3	0.5	0.1
	75	26	25	53	1	0.7	0.5
	80	22	20	36	0.8	0.5	0.3
	85	19	15	30	0.1	0.3	0.08
	90	17	10	19	0.5	0.3	0.2
larger than 10K	65	31	35	57	0.7	0.7	0.1
	70	22	30	50	0.3	0.1	0.04
	75	18	25	45	0.1	0.2	0.08
	80	14	20	37	0.03	0.4	0.3
	85	13	15	33	0.7	0.09	0.2
	90	12	10	27	0.03	0.8	0.5
	includes municipalities without CCT data as 0% coverage (low group)						

6. Concluding Remarks

- Supporting previous studies, variables related to **inequality, literacy rates, previous crime levels, institutional capabilities, conflict and drug-trafficking** were reported as significant determinants of homicide rates.
- By 2018, **the average homicide rates were lower for high CCT coverage municipalities** when compared to synthetic copies made out of a pool of low CCT coverage municipalities.

Implications

- CCT programs appear to be comprehensive policies as they can tackle multiple issues such as **poverty, low education outcomes and violence**.
- Given funding constraints, it seems that investing in the expansion of this policy in the Pacific region can be an effective way to improve developmental outcomes in several areas.
- The framework of this paper can be considered a **data science framework to test the impact of regional policies**

BMA → determinants of Y → Synthetic Controls → The impact of X on Y

6. Further research

- The robustness of the result can be tested by changing the thresholds for low and high CCT coverage. Instead of 30% and 70%, one can use thresholds such as **20%-80% and 25%-75%**.
- A dataset of determinants of crime based on previous literature (instead of the determinants found with the BMA) can be assembled. **This new dataset can be used as the input for the synthetic control analysis.**
- **How can we integrate spatial effects and Synthetic Controls?** Spatial filtering? Adding a distance indicator as one of the determinants in the Synthetic Controls framework?

7. Open Science

"The current state of Science in terms of transparency and openness is prompting for action... the term "reproducibility" is also gaining traction. ... its definition alludes to the need of scientific results to be accompanied by enough information and detail so they could be repeated by a third party" (Rey et al. 2021)

How can you replicate these results?

Synthetic Controls

<https://github.com/jfsantosm/2021a-crime-ccts-colombia>

Click on [launch binder]

BMS [forthcoming]

Do you like R and R studio?

You can literally replicate these slides here

<https://github.com/Chair-International-Economics/BBS-slides-2021-july>

Thank you very much for your attention

personal website: <https://felipe-santos.rbind.io>

slides available at: <https://bbs-felipe.netlify.app/>

:) Working paper available on request felipe.santos@tu-dresden.de :)



Quantitative Regional and
Computational Science lab

<https://quarcs-lab.org>



Chair of International Economics
Technische Universität Dresden
Germany

<https://tu-dresden.de/bu/wirtschaft/vwl/iwb>