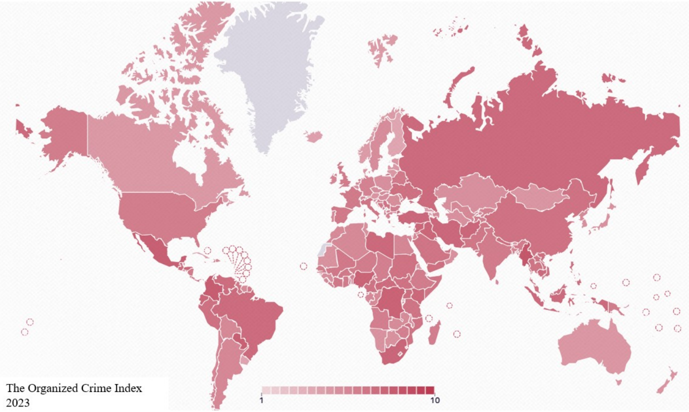
> April 2024

Mafia, Politics & Machine Predictions

Gian Maria Campedelli[1], **Gianmarco Daniele**[2,3], **Marco Le Moglie**[4]
[1] Fondazione Bruno Kessler
[2] University of Milan
[3] Bocconi University
[4] Università Cattolica

Motivation

- 80% of the global population lives in countries where organized crime (OC) presents a high risk [Organized Crime Index, 2023], with negative effects on society, e.g. economic growth, human capital, influence on **politics** [Alesina et al., 2019; Daniele and Dipoppa, 2017; Pinotti, 2015b; Sviatschi, 2022]



Problem: Detection and Measurement

- **It is difficult to fight OC because detection is imperfect and costly** [Pinotti, 2015a; United Nations Office on Drugs and Crime, 2016]
  - Illegal phenomenon
  - OC activities tend to be secretive: corruption, money laundering, and obstruction of justice

Problem: Detection and Measurement

- **It is difficult to fight OC because detection is imperfect and costly** [Pinotti, 2015a; United Nations Office on Drugs and Crime, 2016]
    - Illegal phenomenon
    - OC activities tend to be secretive: corruption, money laundering, and obstruction of justice
- In turn, OC measurement is imperfect
    - Researchers resort to detection measures (e.g. drug seizures) or visible signs of their presence (e.g. homicides)
- These issues are more complex when we deal with OC infiltration into politics, as criminals mostly rely on corruption making detection more challenging

Problem: Detection and Measurement

- **It is difficult to fight OC because detection is imperfect and costly** [Pinotti, 2015a; United Nations Office on Drugs and Crime, 2016]
    - Illegal phenomenon
    - OC activities tend to be secretive: corruption, money laundering, and obstruction of justice
- In turn, OC measurement is imperfect
    - Researchers resort to detection measures (e.g. drug seizures) or visible signs of their presence (e.g. homicides)
- These issues are more complex when we deal with OC infiltration into politics, as criminals mostly rely on corruption making detection more challenging

Can we use machine learning (ML) to create a measure of OC presence in politics?

Problem: Detection and Measurement

- **It is difficult to fight OC because detection is imperfect and costly** [Pinotti, 2015a; United Nations Office on Drugs and Crime, 2016]
    - Illegal phenomenon
    - OC activities tend to be secretive: corruption, money laundering, and obstruction of justice
- In turn, OC measurement is imperfect
    - Researchers resort to detection measures (e.g. drug seizures) or visible signs of their presence (e.g. homicides)
- These issues are more complex when we deal with OC infiltration into politics, as criminals mostly rely on corruption making detection more challenging

Can we use machine learning (ML) to create a measure of OC presence in politics?

Can we use this new measure to facilitate detection and study OC influence on politics?

What we do: focus on OC and Politics

- By leveraging ML algorithms, we predict local governments in Italy with a high risk of mafia infiltration
    - We create a synthetic measure of mafia infiltration in politics based on city council dismissals for mafia infiltration by the national government
    - We propose this indicator as a tool to improve the detection of mafia infiltration in local politics

What we do: focus on OC and Politics

- By leveraging ML algorithms, we predict local governments in Italy with a high risk of mafia infiltration
  - We create a synthetic measure of mafia infiltration in politics based on city council dismissals for mafia infiltration by the national government
  - We propose this indicator as a tool to improve the detection of mafia infiltration in local politics
- A stronger state presence... Does a public spending shock discourage or promote OC influence on local politics?
  - An increase in public investments might promote economic growth, reducing the grip of OC
  - More spending might attract OC into politics

Preview

- ML to predict mafia infiltration in local politics measured by city councils dismissals for mafia infiltration:

Preview

- ML to predict mafia infiltration in local politics measured by city councils dismissals for mafia infiltration:
  - The index - based only on electoral and budget variables - predicts up to 96% of municipalities infiltrated by mafias – also using data at $t-1$ and $t-2$

Preview

- ML to predict mafia infiltration in local politics measured by city councils dismissals for mafia infiltration:
    - The index - based only on electoral and budget variables - predicts up to 96% of municipalities infiltrated by mafias – also using data at $t-1$ and $t-2$
    - Focusing on the top 4.5% (1%) of the predicted risk in a given year, that is 360 (72) municipalities, we ensure recall that is on average around 89% (50%), with the true positives on average being around 15% (28%).

Preview

- ML to predict mafia infiltration in local politics measured by city councils dismissals for mafia infiltration:
    - The index - based only on electoral and budget variables - predicts up to 96% of municipalities infiltrated by mafias – also using data at $t-1$ and $t-2$
    - Focusing on the top 4.5% (1%) of the predicted risk in a given year, that is 360 (72) municipalities, we ensure recall that is on average around 89% (50%), with the true positives on average being around 15% (28%).
    - While false positives are generally a "bad thing", here they might represent municipalities suitable for targeted investigations

Preview

- ML to predict mafia infiltration in local politics measured by city councils dismissals for mafia infiltration:
    - The index - based only on electoral and budget variables - predicts up to 96% of municipalities infiltrated by mafias – also using data at $t-1$ and $t-2$
    - Focusing on the top 4.5% (1%) of the predicted risk in a given year, that is 360 (72) municipalities, we ensure recall that is on average around 89% (50%), with the true positives on average being around 15% (28%).
    - While false positives are generally a "bad thing", here they might represent municipalities suitable for targeted investigations
- We then test a new research question that would be difficult to explore without such a fine-grained measure:

Preview

- ML to predict mafia infiltration in local politics measured by city councils dismissals for mafia infiltration:
  - The index - based only on electoral and budget variables - predicts up to 96% of municipalities infiltrated by mafias – also using data at $t-1$ and $t-2$
  - Focusing on the top 4.5% (1%) of the predicted risk in a given year, that is 360 (72) municipalities, we ensure recall that is on average around 89% (50%), with the true positives on average being around 15% (28%).
  - While false positives are generally a "bad thing", here they might represent municipalities suitable for targeted investigations
- We then test a new research question that would be difficult to explore without such a fine-grained measure:
  - Does a positive shock in public spending targeting areas affected by OC (the 2007-2013 EU funds wave) reduce mafia infiltration in politics?

Preview

- ML to predict mafia infiltration in local politics measured by city councils dismissals for mafia infiltration:
    - The index - based only on electoral and budget variables - predicts up to 96% of municipalities infiltrated by mafias – also using data at $t-1$ and $t-2$
    - Focusing on the top 4.5% (1%) of the predicted risk in a given year, that is 360 (72) municipalities, we ensure recall that is on average around 89% (50%), with the true positives on average being around 15% (28%).
    - While false positives are generally a "bad thing", here they might represent municipalities suitable for targeted investigations
- We then test a new research question that would be difficult to explore without such a fine-grained measure:
    - Does a positive shock in public spending targeting areas affected by OC (the 2007-2013 EU funds wave) reduce mafia infiltration in politics? **NO, it increases it.**

Literature

Literature

- Consequences of OC on politics [Acemoglu et al., 2013; Alesina et al., 2019; Daniele and Dipoppa, 2017]

Literature

- Consequences of OC on politics [Acemoglu et al., 2013; Alesina et al., 2019; Daniele and Dipoppa, 2017]
- Anti-OC policy:
  - Unintended effects of anti-OC policies [Battiston et al., 2022; Castillo and Kronick, 2020; Daniele and Dipoppa, 2023; Dell, 2015; Lessing, 2017]
  - City councils dismissals for mafia [Acconcia et al., 2014; Daniele and Geys, 2015; Fenizia, 2018; Galletta, 2017]

Literature

- Consequences of OC on politics [Acemoglu et al., 2013; Alesina et al., 2019; Daniele and Dipoppa, 2017]
- Anti-OC policy:
  - Unintended effects of anti-OC policies [Battiston et al., 2022; Castillo and Kronick, 2020; Daniele and Dipoppa, 2023; Dell, 2015; Lessing, 2017]
  - City councils dismissals for mafia [Acconcia et al., 2014; Daniele and Geys, 2015; Fenizia, 2018; Galletta, 2017]
- Redistributive policy (EU funds) and economic development [Alesina and Perotti, 2002; Becker et al., 2010a]
- Rapacity effect [Dube and Vargas, 2013]
- Machine learning in economics [Ash et al., 2020; Athey, 2018; Glaeser et al., 2016; Kleinberg et al., 2018; Mohler et al., 2015]

Part I:
Mafia Infiltration and Machine Predictions.

City council dismissals

Since 1991, the Italian authorities can dismiss a city council if there is evidence of mafia infiltration, specifically:

- Direct/indirect contacts of local politicians with organized crime groups
- Ability to influence the decision-making process of local politicians (e.g. directing public procurement towards criminal firms, hiring decisions, building permits, etc.)

City council dismissals

Since 1991, the Italian authorities can dismiss a city council if there is evidence of mafia infiltration, specifically:

- Direct/indirect contacts of local politicians with organized crime groups
- Ability to influence the decision-making process of local politicians (e.g. directing public procurement towards criminal firms, hiring decisions, building permits, etc.)

Start of the investigation
Based on a clue, phone calls traced

Investigation
6-months – lead by national police

Dismissal of the city council
By the Ministry of Interior

City council dismissals

Since 1991, the Italian authorities can dismiss a city council if there is evidence of mafia infiltration, specifically:

- Direct/indirect contacts of local politicians with organized crime groups
- Ability to influence the decision-making process of local politicians (e.g. directing public procurement towards criminal firms, hiring decisions, building permits, etc.)

Start of the investigation
Based on a clue, phone calls traced

Investigation
6-months – lead by national police

Dismissal of the city council
By the Ministry of Interior

After the dismissal, three appointed bureaucrats rule the municipality up to 24 months

City council dismissals

Since 1991, the Italian authorities can dismiss a city council if there is evidence of mafia infiltration, specifically:

- Direct/indirect contacts of local politicians with organized crime groups
- Ability to influence the decision-making process of local politicians (e.g. directing public procurement towards criminal firms, hiring decisions, building permits, etc.)

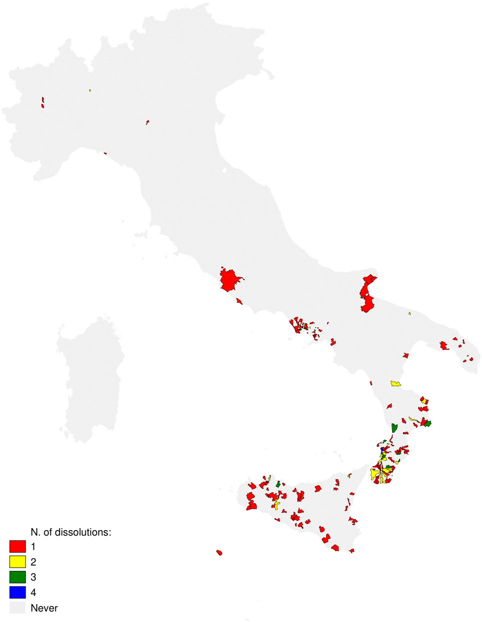
Start of the investigation
Based on a clue, phone calls traced

Investigation
6-months – lead by national police

Dismissal of the city council
By the Ministry of Interior

After the dismissal, three appointed bureaucrats rule the municipality up to 24 months Most of the 379 dismissals took place in three Southern regions: Calabria, Campania and Sicily

Geographical distribution of dismissals



Predictive Design

- Predicting **detected** mafia infiltration (yearly, municipal-level)
- Y is constructed as follows:
  - 1 if a city council was dissolved during the mandate of the mayor
  - 0 otherwise

▸ Dissolved Councils and Infiltration distribution over time

Predictive Design

- Predicting **detected** mafia infiltration (yearly, municipal-level)
- Y is constructed as follows:
    - 1 if a city council was dissolved during the mandate of the mayor
    - 0 otherwise ▸ Dissolved Councils and Infiltration distribution over time
- Observations at the municipality-year level (N=152k)
    - Training the Model: Use a random portion of the data to train the model
    - Cross-validation of the model
    - Stratified sampling of X and Y (hence=no temporal order preserved)

Predictive Design

- Predicting **detected** mafia infiltration (yearly, municipal-level)
- Y is constructed as follows:
  - 1 if a city council was dissolved during the mandate of the mayor
  - 0 otherwise ▸ Dissolved Councils and Infiltration distribution over time
- Observations at the municipality-year level (N=152k)
  - Training the Model: Use a random portion of the data to train the model
  - Cross-validation of the model
  - Stratified sampling of X and Y (hence=no temporal order preserved)
- Nine algorithmic approaches (LR, Lasso, Ridge, Elastic Net, DT, RF, GBoosting, XGBoost, DNN)
  - Optimization via Hyperparameter Grid Search (1500+ model candidates, 200 tested)
  - Run-time: 20 hours on 8-core machine, $\sim$ 5 hours on 32-core cluster

Predictive Design

- Predicting **detected** mafia infiltration (yearly, municipal-level)
- Y is constructed as follows:
  - 1 if a city council was dissolved during the mandate of the mayor
  - 0 otherwise ▸ Dissolved Councils and Infiltration distribution over time
- Observations at the municipality-year level (N=152k)
  - Training the Model: Use a random portion of the data to train the model
  - Cross-validation of the model
  - Stratified sampling of X and Y (hence=no temporal order preserved)
- Nine algorithmic approaches (LR, Lasso, Ridge, Elastic Net, DT, RF, GBoosting, XGBoost, DNN)
  - Optimization via Hyperparameter Grid Search (1500+ model candidates, 200 tested)
  - Run-time: 20 hours on 8-core machine, $\sim$ 5 hours on 32-core cluster
- A municipal-year observation is infiltrated if the prediction is $>$0.5

Predictive Design

- Predicting **detected** mafia infiltration (yearly, municipal-level)
- Y is constructed as follows:
    - 1 if a city council was dissolved during the mandate of the mayor
    - 0 otherwise ▸ Dissolved Councils and Infiltration distribution over time
- Observations at the municipality-year level (N=152k)
    - Training the Model: Use a random portion of the data to train the model
    - Cross-validation of the model
    - Stratified sampling of X and Y (hence=no temporal order preserved)
- Nine algorithmic approaches (LR, Lasso, Ridge, Elastic Net, DT, RF, GBoosting, XGBoost, DNN)
    - Optimization via Hyperparameter Grid Search (1500+ model candidates, 200 tested)
    - Run-time: 20 hours on 8-core machine, $\sim$ 5 hours on 32-core cluster
- A municipal-year observation is infiltrated if the prediction is >0.5
- Prediction phase/optimization $\longrightarrow$ Explainability
    - SHAP: Global and local interpretability of ML models ▸ Details

Machine Learning Approach

- **Chosen metric: Recall** (i.e., maximize true positives or penalize false negatives)

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

- Alternative metric: Precision

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

A Highly Unbalanced Problem

Sampling	Y=0	Y=1	% (Y=1)
Original	105,596	887	0.83
SMOTE	105,596	105,596	50.00
ADASYN	105,582	105,596	49.99
SMOTE+Tomek	105,595	105,595	50.00

- Challenging prediction application: highly unbalanced distribution → synthetic oversampling in the training set
- In the test set the distribution remains identical (Y=1 is 0.83% of the total observations)

▸ Details ▸ Infiltration distribution over time

Feature Space

- Time range: 2001-2020 (quasi-universe of municipalities in Italy)
- 200+ variables per each observation
    - Regional variables

Feature Space

- Time range: 2001-2020 (quasi-universe of municipalities in Italy)
- 200+ variables per each observation
  - Regional variables
  - Electoral variables (ideology, number of competitors, mayor's demographics, etc)

▸ List

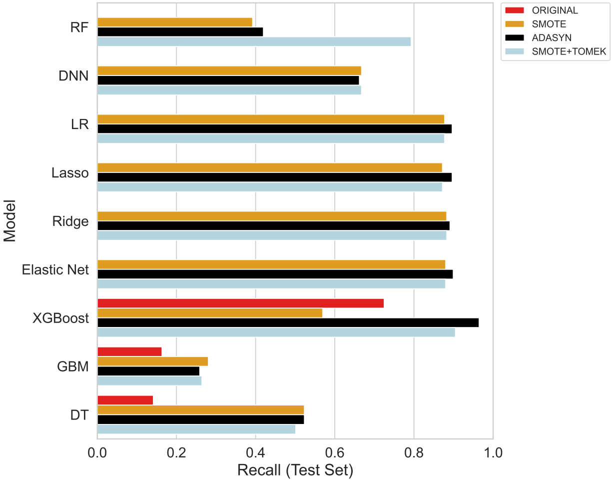
Feature Space

- Time range: 2001-2020 (quasi-universe of municipalities in Italy)
- 200+ variables per each observation
  - Regional variables
  - Electoral variables (ideology, number of competitors, mayor's demographics, etc)
    ▸ List
  - 22 Public Spending variables (at t and t-1), e.g. Local Police, Education, Environment, Health, Tourism

Feature Space

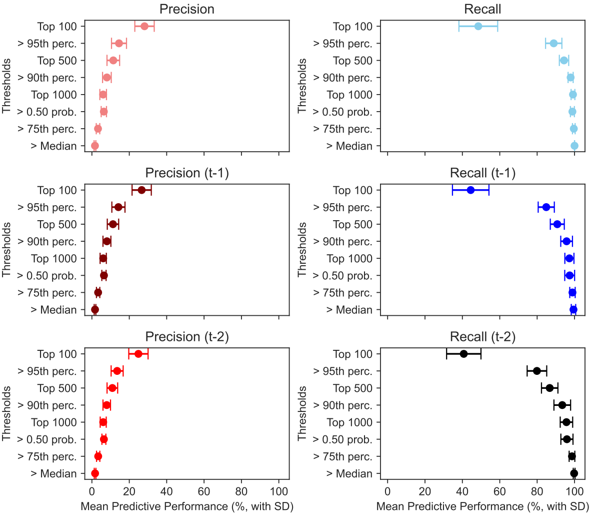
- Time range: 2001-2020 (quasi-universe of municipalities in Italy)
- 200+ variables per each observation
  - Regional variables
  - Electoral variables (ideology, number of competitors, mayor's demographics, etc) ▸ List
  - 22 Public Spending variables (at t and t-1), e.g. Local Police, Education, Environment, Health, Tourism
  - We differentiate across different versions of the spending variables (e.g. current and capital spending) ▸ Details

Prediction Performance: Recall



▸ Revoked Vs Confirmed dismissals

Recall and Precision for Various Rankings, by Year



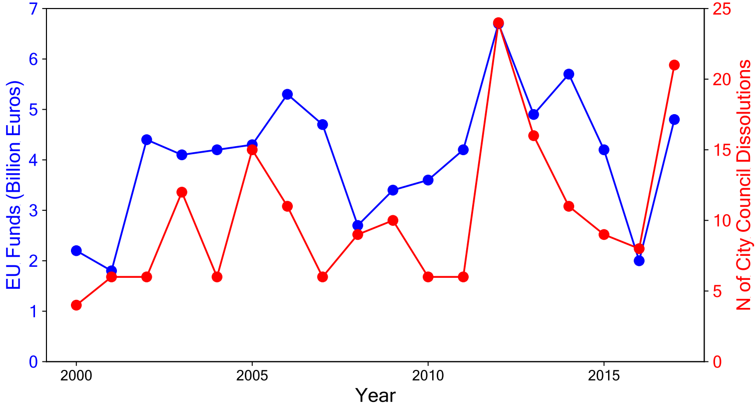
▸ By macro-regions ▸ Where we contribute

Additional Results

- Which are the most relevant predictors? ▸ Global SHAP ▸ Local SHAP ▸ Example
- Variation across time and space ▸ Map
- Variable distribution ▸ Graph ▸ Graph across macro-regions
- External Validity ▸ 2021-2023 Prediction
- Variation around the dismissal ▸ Prediction dynamics
- Correlations with other crimes ▸ Correlations
- Additional ML exercises ▸ Table
- Random Seeds ▸ Graph

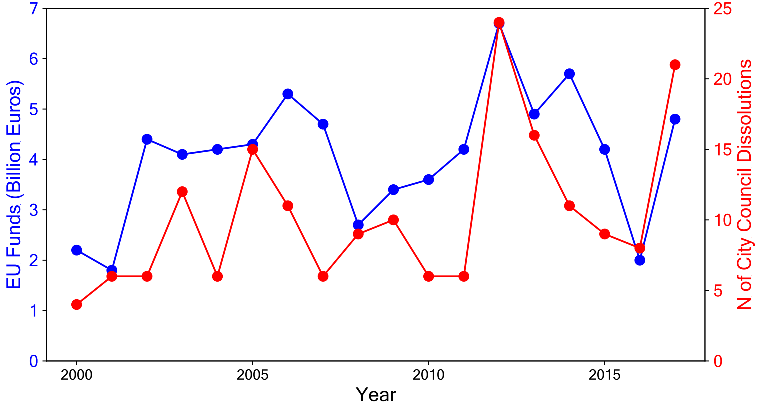
Part II: Redistributive policy and on mafia infiltration in politics

A stronger State presence



- Does a positive shock in public spending affect mafia infiltration in politics?
  - Transfers can either foster economic growth, reducing the grip of mafia, or they might push mafia toward local governments managing new funds

A stronger State presence



- Does a positive shock in public spending affect mafia infiltration in politics?
  - Transfers can either foster economic growth, reducing the grip of mafia, or they might push mafia toward local governments managing new funds
- Why do transfers promote growth in some areas and not in others?
  - Becker et al [2010b] show EU funds worked in most areas but limited effect in Southern Italy

Transfers windfall and Mafia

We study if and how a transfers windfall affects mafia presence in local politics:

- The 2007-2013 EU funds
- Funds are disproportionally allocated to "convergence" regions, i.e. regions with a GDP below 75% EU average - Southern regions
- Budget increase from about 30 to 56 billion Euros

Transfers windfall and Mafia

We study if and how a transfers windfall affects mafia presence in local politics:

- The 2007-2013 EU funds
- Funds are disproportionally allocated to "convergence" regions, i.e. regions with a GDP below 75% EU average - Southern regions
- Budget increase from about 30 to 56 billion Euros

We target the impact of the 2007-2013 wave which may have affected the risk of mafia infiltration via two main channels:

- Larger budget (i.e., more economic resources assigned) ▸ Graph
- Increased decentralization in spending decision (local institutions are key spending authority) & higher flexibility in implementation/simplification of rules
- Municipalities can invest more in capital spending and provide more subsidies to local firms/NGOs

Identification

Identification: Diff-in-Disc

$$\begin{aligned} \text{Infiltration}_{i,t} = {} & \alpha + \beta \text{Treated}_i \times \text{Post2006}_t + \eta f(\text{Distance}_i) \times \text{Post2006}_t \\ & + \gamma \text{Treated}_i \times f(\text{Distance}_i) \times \text{Post2006}_t + FE_i + FE_t + \varepsilon_{i,t}, \end{aligned} \tag{1}$$

- Infiltration$_{i,t}$= Infiltration risk for municipality *i* in year *t*; period 2001-2013, treatment since 2007
- We consider municipalities in 4 regions: Lazio and Molise for center Italy, and Campania and Puglia for southern Italy: Treated$_i$, i.e., being a southern municipality
- f(Distance), i.e., a local polynomial of the distance in Km from the border defined for different bandwidths: 5, 10, 25, 50, and 100 km and the entire region. Either linear or quadratic specification
- We use a triangular kernel to weigh observations
- $FE_i + FE_t$ are municipality and year fixed effects
- $\varepsilon_{i,t}$ standard errors are either clustered at the municipal level or computed by bootstrapping

Identification: Diff-in-Disc

The spatial diff-in-disc [Butts, 2023] has three main advantages in our context, as it reduces...:

- ...the effect of any potential unbalancing around the administrative discontinuity in time-fixed/slowly changing characteristics

Identification: Diff-in-Disc

The spatial diff-in-disc [2023] has three main advantages in our context, as it reduces...:

- ...the effect of any potential unbalancing around the administrative discontinuity in time-fixed/slowly changing characteristics
- ...the confounding effect of other compound treatments as long as the latter do not vary simultaneously to the treatment

Identification: Diff-in-Disc

The spatial diff-in-disc [2023] has three main advantages in our context, as it reduces...:

- ...the effect of any potential unbalancing around the administrative discontinuity in time-fixed/slowly changing characteristics
- ...the confounding effect of other compound treatments as long as the latter do not vary simultaneously to the treatment
- ...the non-classical measurement error induced by the non-random training of the prediction measure (time-invariant part), while the comparability of treated and control units reduces concerns related to its time-varying part

Identification: Diff-in-Disc

The spatial diff-in-disc [2023] has three main advantages in our context, as it reduces...:

- ...the effect of any potential unbalancing around the administrative discontinuity in time-fixed/slowly changing characteristics
- ...the confounding effect of other compound treatments as long as the latter do not vary simultaneously to the treatment
- ...the non-classical measurement error induced by the non-random training of the prediction measure (time-invariant part), while the comparability of treated and control units reduces concerns related to its time-varying part

Three necessary conditions/assumptions:

- Non-sorting around the threshold in both the pre-and post-treatment

Identification: Diff-in-Disc

The spatial diff-in-disc [2023] has three main advantages in our context, as it reduces...:

- ...the effect of any potential unbalancing around the administrative discontinuity in time-fixed/slowly changing characteristics
- ...the confounding effect of other compound treatments as long as the latter do not vary simultaneously to the treatment
- ...the non-classical measurement error induced by the non-random training of the prediction measure (time-invariant part), while the comparability of treated and control units reduces concerns related to its time-varying part

Three necessary conditions/assumptions:

- Non-sorting around the threshold in both the pre-and post-treatment
- Parallel trend assumption between treated/control municipalities

Identification: Diff-in-Disc

The spatial diff-in-disc [2023] has three main advantages in our context, as it reduces...:

- ...the effect of any potential unbalancing around the administrative discontinuity in time-fixed/slowly changing characteristics
- ...the confounding effect of other compound treatments as long as the latter do not vary simultaneously to the treatment
- ...the non-classical measurement error induced by the non-random training of the prediction measure (time-invariant part), while the comparability of treated and control units reduces concerns related to its time-varying part

Three necessary conditions/assumptions:

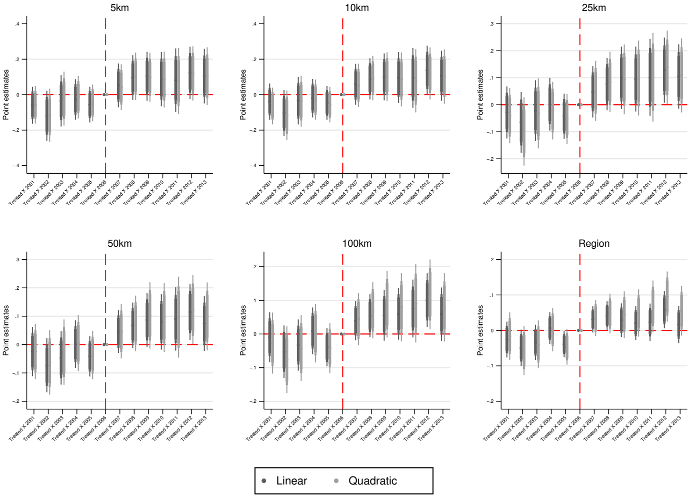
- Non-sorting around the threshold in both the pre-and post-treatment
- Parallel trend assumption between treated/control municipalities
- Changes in covariates induced by the treatment balance between T and C

▸ Balancing

Results: Diff-in-Disc

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.
Treat X Post 2006	0.142***	0.132***	0.138***	0.141***	0.120***	0.138***	0.111***	0.115***	0.084***	0.112***	0.055***	0.083***
	(0.040)	(0.041)	(0.038)	(0.040)	(0.033)	(0.037)	(0.026)	(0.033)	(0.020)	(0.027)	(0.012)	(0.017)
Observations	1,298	1,298	2,023	2,023	3,934	3,934	7,105	7,105	10,904	10,904	17,111	17,111
R-squared	0.701	0.706	0.700	0.700	0.723	0.724	0.784	0.784	0.822	0.823	0.806	0.806
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	2nd	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.0757	.0887	.0646	.0648	.0471	.0502	.0452	.0453	.0438	.0455	.0482	.0495
Bootstrap	.0006	.0018	.0008	.001	.0002	.0004	0	.0009	.0001	.0001	0	0
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	.17	.17	.15	.15	.12	.12	.11	.11	.11	.11	.1	.1

Results: Event Study



Additional Tests

- Placebo with fake borders Figure
- Robust coefficient Table Figure
- Random Seeds Table
- Placebo: EU Funds but no risk of Mafia Infiltration Table
- Removing capital spending from the prediction Event Study
- Removing 2007-2013 from the training set Table
- Mafia-related crimes Table Event Study Table
- Mechanisms:
  - Lazio Vs Campania (only decentralization) Table
  - Puglia Vs Molise (decentralization + funds increase) Table

Conclusions

- Our ML-based approach correctly predicts 96% of infiltrated municipalities with a good trade-off between recall and precision

Conclusions

- Our ML-based approach correctly predicts 96% of infiltrated municipalities with a good trade-off between recall and precision
- Targeting a highly ranked subset of predictions allows linking recall and precision meaningfully

Conclusions

- Our ML-based approach correctly predicts 96% of infiltrated municipalities with a good trade-off between recall and precision
- Targeting a highly ranked subset of predictions allows linking recall and precision meaningfully

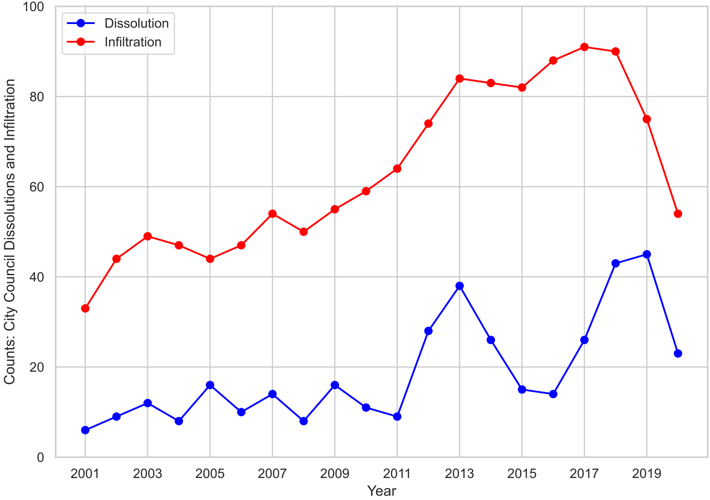
- The 2007-2013 wave of EU funds increased the risk of mafia infiltration in treated municipalities
- The mix of predictive and causal methods offers insights into the effects of increased state presence (transfers) in OC affected areas

> Thanks for attending!

Questions? Comments?

Appendix.

Dissolved Councils and Infiltration distribution over time



▸ Back

ML Explainability via SHAP

- SHAP reveals the most influential features in a model’s prediction output
- It compares the model’s output when a specific feature is included versus when it is excluded
- SHAP provides both local and global measures for each feature, helping us assess its impact on the model’s decisions

▸ Back

SMOTE (Synthetic Minority Over-sampling Technique)

Mathematical Formulation

- Let A be a minority class instance, and B and C be its k-nearest neighbors.
- For each A, generate synthetic instances A' by connecting A with some of its neighbors B or C in the feature space.
- The synthetic instance A' is given by $A' = A + \lambda \times (B - A)$, where λ is a random value between 0 and 1.

▸ Back

ADASYN (Adaptive Synthetic Sampling)

Mathematical Formulation

- Incorporates a density distribution factor to adaptively generate synthetic instances.
- For each minority class instance, calculate the number of synthetic instances to generate based on the density ratio.
- Use the same formula as SMOTE to generate synthetic instances, but with an adjusted λ based on the density distribution.
- Promotes the creation of more synthetic instances for minority instances in denser regions.

▸ Back

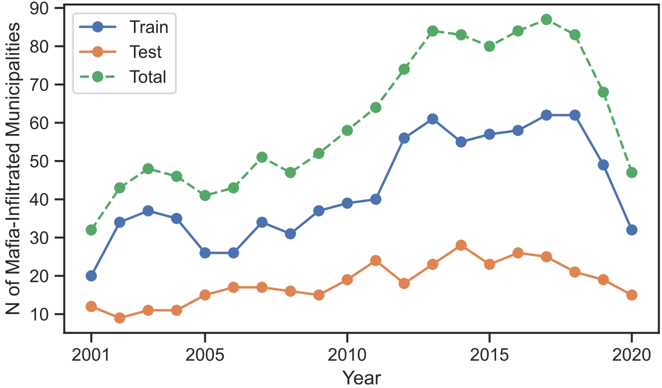
SMOTE-Tomek (Combining SMOTE with Tomek Links)

Mathematical Formulation

- Identify Tomek links, pairs of instances (A, B) where A is the nearest neighbor of B but they belong to different classes.
- Apply SMOTE only to instances involved in Tomek links, focusing on generating synthetic instances for instances that contribute to noise.
- After SMOTE, remove Tomek links to clean the dataset from noisy and irrelevant synthetic instances.

▸ Back

Infiltration distribution over time



Back

Political Variables

For each municipality election, we have information on:

- Ideology/political placement of winning coalition/party
- Share of votes for winning candidate
- Number of competitors
- Sex of the mayor
- Educational background of the mayor (college degree or not)
- Incumbency
- Birth Location (municipality i or not)

▸ Back

Public Spending Variables/2

For each typology, we construct four variables. Given spending typology x:

- Share of *current* expenses

$$\text{Share Curr. } (x)_{i,t} = \frac{\text{Curr}(x)_{i,t}}{\sum_{x \in X} \text{Curr}(x)_{i,t}} \tag{2}$$

- Share of *capital* expenses

$$\text{Share Cap. } (x)_{i,t} = \frac{\text{Cap}(x)_{i,t}}{\sum_{x \in X} \text{Cap}(x)_{i,t}} \tag{3}$$

- Expense rate

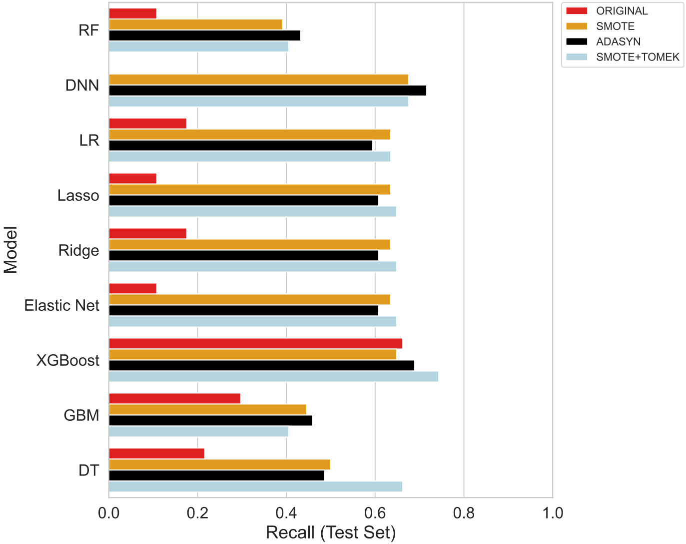
$$\text{Rate}(x)_{i,t} = \left(\frac{\text{Curr}(x)_{i,t} + \text{Cap}(x)_{i,t}}{\text{pop}_{i,t}} \right) \times 10k \tag{4}$$

- Share of *current* expenses out of total expenses:

$$\text{Share Curr. Global } (x)_{i,t} = \frac{\text{Curr}(x)_{i,t}}{\sum_{x \in X} (\text{Curr}(x)_{i,t} + \text{Cap}(x)_{i,t})} \tag{5}$$

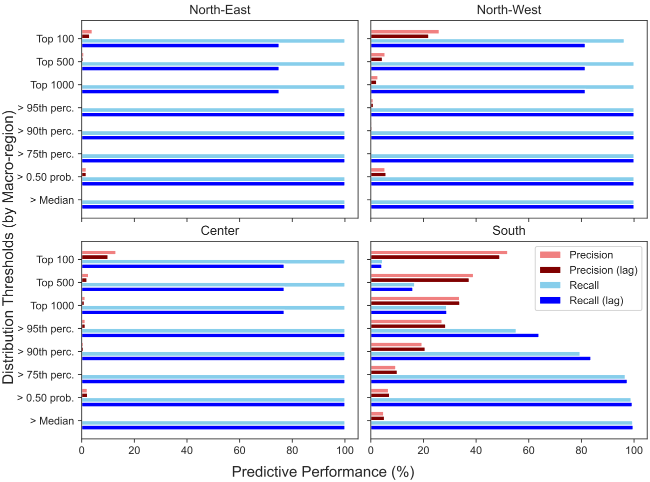
▸ Back

Predictive Performance



back

Recall and Precision for Various Rankings, by Macro-Region



▸ Back

Investigations step

- We can improve the first steps of the detection process, i.e. the probability of starting an investigation (“proactive” instead of “reactive” investigations)

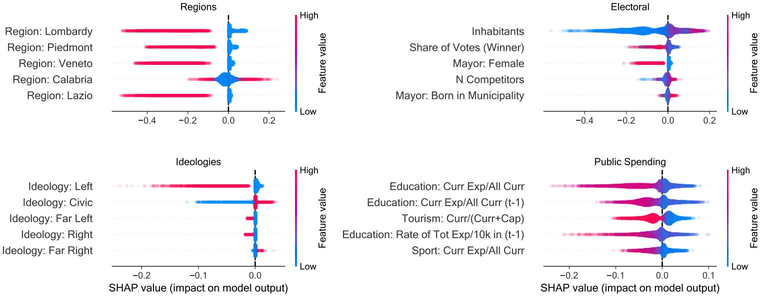
Start of the investigation
Based on a clue, phone calls traced

Investigation
6-months – lead by national police

Dismissal of the city council
By the Ministry of Interior

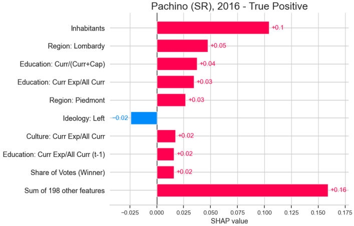
Back

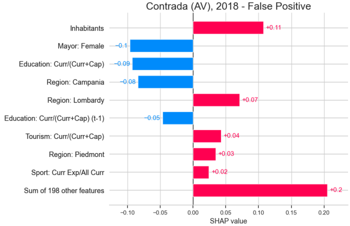
SHAP: Explainability Results

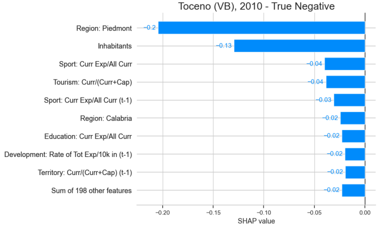


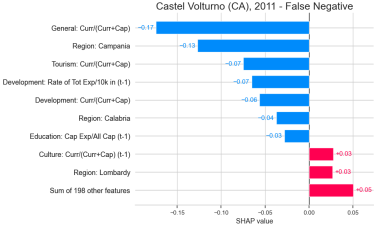
Back

Locally Explainable Predictions


Pachino (SR), 2016 - True Positive

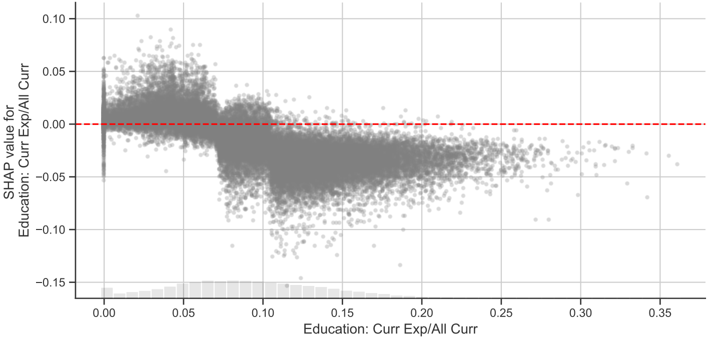

Contrada (AV), 2018 - False Positive


Toceno (VB), 2010 - True Negative


Castel Volturno (CA), 2011 - False Negative

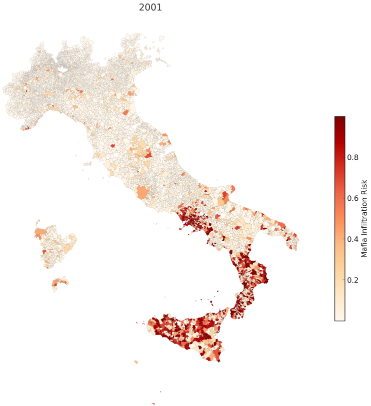
▶ Back

Feature Explainability - Example



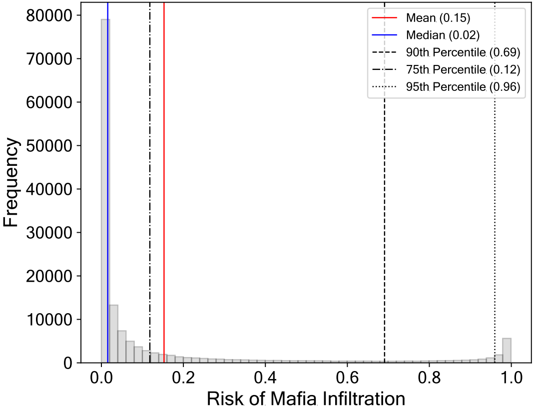
▸ Back

Predicted Infiltration Risk 2001-2020, Distribution in Time and Space



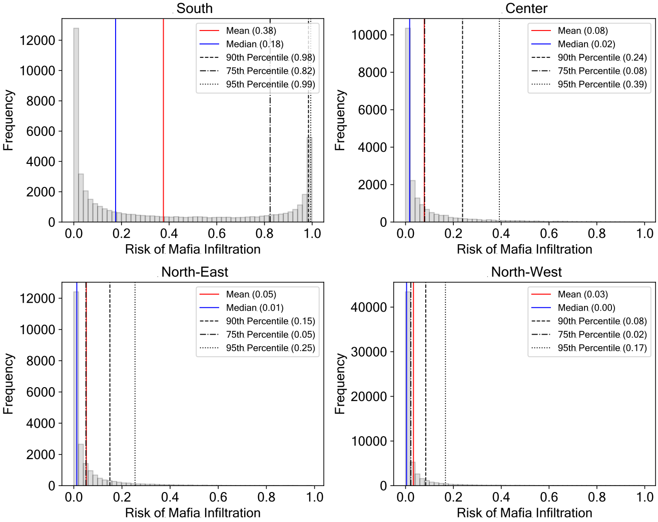
Back

Predicted Infiltration Risk 2001-2020, Distribution



▸ Back

Predicted Infiltration Risk 2001-2020, Distribution (by Macro-Region)



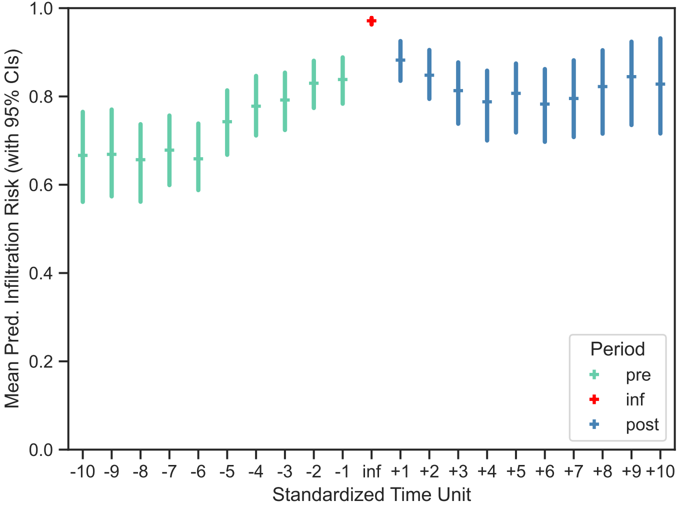
▶ Back

External Validity: Municipalities Dissolved 2021-2023

Municipality	Region	dismissal year	2020	Avg. 2016-2020	Last term
Squinzano	Apulia	30/01/21	98%	92%	97%
Guardavalle	Calabria	23/02/21	100%	100%	100%
Carovigno	Apulia	12/03/21	76%	76%	82%
Barrafranca	Sicily	16/04/21		100%	
Marano di Napoli	Campania	18/06/21	100%	100%	100%
San Giuseppe Jato	Sicily	09/07/21	81%	53%	42%
Villaricca	Campania	06/08/21	91%	82%	82%
Foggia	Apulia	06/08/21	20%	31%	22%
Nocera Terinese	Calabria	30/08/21	100%	95%	99%
Simeri Crichi	Calabria	30/08/21	99%	99%	99%
Rosarno	Calabria	30/08/21	100%	98%	98%
Calatabiano	Sicily	30/08/21	97%	98%	98%
Bolognetta	Sicily	18/11/21		99%	
Ostuni	Apulia	27/12/21	42%	32%	24%
Castellammare di Stabia	Campania	24/02/22	90%	75%	94%
Trinitapoli	Apulia	05/04/22	65%	51%	65%
Torre Annunziata	Campania	06/05/22	20%	67%	69%
Portigliola	Calabria	22/05/22	0%	5%	5%
San Giuseppe Vesuviano	Campania	09/06/22	100%	92%	99%
Soriano Calabro	Calabria	17/06/22	100%	92%	99%
Neviano	Apulia	05/08/22	53%	41%	53%
Cosoleto	Calabria	21/11/22	33%	12%	13%
Nettuno	Lazio	21/11/22	36%	54%	50%
Anzio	Lazio	21/11/22	81%	71%	80%
Sparanise	Campania	19/12/22	98%	97%	97%
Scilla	Calabria	11/04/23	97%	99%	97%
Castiglione di Sicilia	Sicily	23/05/23	100%	79%	75%
Rende	Calabria	28/06/23	100%	99%	100%
Orta Nova	Apulia	18/07/23	93%	73%	93%
Palagonia	Sicily	09/08/23	93%	93%	92%
Acquaro	Calabria	18/09/23	87%	95%	87%
Caivano	Campania	17/10/23	95%	98%	94%
Capistrano	Calabria	17/10/23	0%	35%	30%

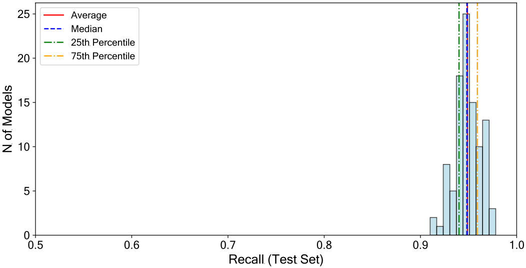
▶ Back

Descriptive Evidence: Predicted Infiltration Risk for Dissolved Municipalities



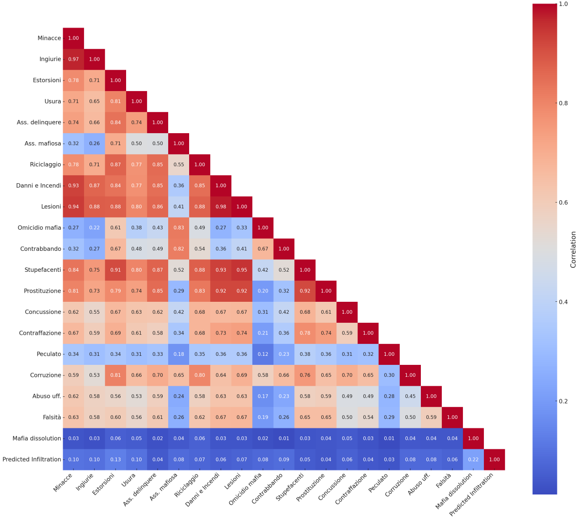
Back

100 random seeds



▶ Back

Predicted Infiltration Risk 2001-2020, Correlations



	Minacce	Ingiurie	Estorsioni	Usura	Ass. delinquere	Ass. mafiosa	Riciclaggio	Danni e Incendi	Lesioni	Omicidio mafia	Contrabbando	Stupefacenti	Prostituzione	Concussione	Contraffazione	Peculato	Corruzione	Abuso uff.	Falsità	Mafia dissolution	Predicted Infiltration
Minacce	1.00																				
Ingiurie	0.97	1.00																			
Estorsioni	0.78	0.71	1.00																		
Usura	0.71	0.65	0.81	1.00																	
Ass. delinquere	0.74	0.66	0.84	0.74	1.00																
Ass. mafiosa	0.32	0.26	0.71	0.50	0.50	1.00															
Riciclaggio	0.78	0.71	0.87	0.77	0.85	0.55	1.00														
Danni e Incendi	0.93	0.87	0.84	0.77	0.85	0.36	0.85	1.00													
Lesioni	0.94	0.88	0.88	0.80	0.86	0.41	0.88	0.98	1.00												
Omicidio mafia	0.27	0.22	0.61	0.38	0.43	0.83	0.49	0.27	0.33	1.00											
Contrabbando	0.32	0.27	0.67	0.48	0.49	0.82	0.54	0.36	0.41	0.67	1.00										
Stupefacenti	0.84	0.75	0.91	0.80	0.87	0.52	0.88	0.93	0.95	0.42	0.52	1.00									
Prostituzione	0.81	0.73	0.79	0.74	0.85	0.29	0.83	0.92	0.92	0.20	0.32	0.92	1.00								
Concussione	0.62	0.55	0.67	0.63	0.62	0.42	0.68	0.67	0.67	0.31	0.42	0.68	0.61	1.00							
Contraffazione	0.67	0.59	0.69	0.61	0.58	0.34	0.68	0.73	0.74	0.21	0.36	0.78	0.74	0.59	1.00						
Peculato	0.34	0.31	0.34	0.31	0.33	0.18	0.35	0.36	0.36	0.12	0.23	0.38	0.36	0.31	0.32	1.00					
Corruzione	0.59	0.53	0.81	0.66	0.70	0.65	0.80	0.64	0.69	0.58	0.66	0.76	0.65	0.70	0.65	0.30	1.00				
Abuso uff.	0.62	0.58	0.56	0.53	0.59	0.24	0.58	0.63	0.63	0.17	0.23	0.58	0.59	0.49	0.49	0.28	0.45	1.00			
Falsità	0.63	0.58	0.60	0.56	0.61	0.26	0.62	0.67	0.67	0.19	0.26	0.65	0.65	0.50	0.54	0.29	0.50	0.59	1.00		
Mafia dissolution	0.03	0.03	0.06	0.05	0.02	0.04	0.06	0.03	0.03	0.02	0.01	0.03	0.04	0.05	0.03	0.01	0.04	0.04	0.04	1.00	
Predicted Infiltration	0.10	0.10	0.13	0.10	0.04	0.08	0.07	0.06	0.07	0.08	0.09	0.05	0.04	0.06	0.04	0.03	0.08	0.08	0.06	0.22	1.00

▶ Back

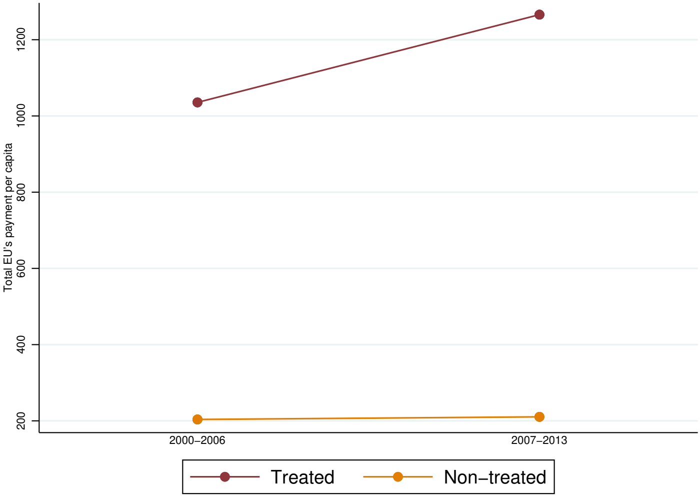
Additional ML exercises

Table: Predictive results of additional ML exercises

Configuration	Obs (% Y=1)	Sampling	Recall	Roc-Auc	Precision
Southern Regions	51,023 (2.31)	Original	0.776	0.841	0.165
		ADASYN	0.903	0.843	0.089
Dissolved Municipalities	3,724 (32.81)	Original	0.932	0.613	0.392
		ADASYN	0.953	0.620	0.395
Dissolved Municipalities (No Yrs After Dissolution)	2,509 (43.76)	Original	0.942	0.697	0.573
		ADASYN	0.955	0.697	0.570
Cross-sectional	7,755 (2.41)	Original	0.357	0.670	0.349
		ADASYN	0.696	0.822	0.246

▸ Back

Dynamics of EU Funds



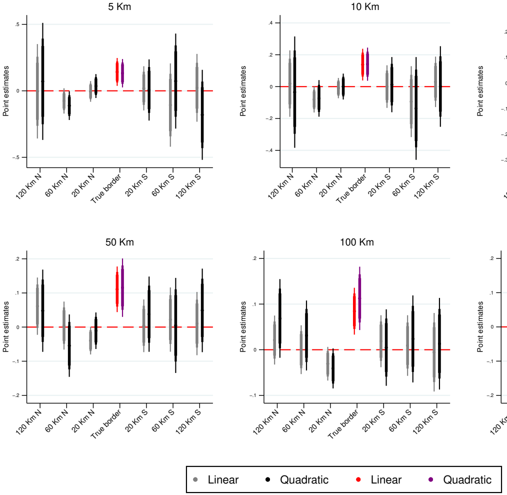
▸ Back

Identification: Balancing

	5 Km		10 Km		25 Km		50 Km		100 Km		Region	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Avg. change population	0.114	0.077	0.109	0.124	0.072	0.132	0.028	0.068	0.023	0.042	-0.098	0.027
Household size	0.006	0.007	0.000	0.005	-0.006	-0.005	-0.007	-0.005	-0.010	-0.010	-0.003	0.003
Aging rate	9.594	9.893	13.268	10.338	15.656	14.563	10.717	15.932	0.730	12.491	-10.216**	-4.822
Poor families	0.018	0.078	-0.110	0.022	-0.199	-0.142	0.027	-0.164	-0.250	0.025	-0.550***	-0.559***
College rate (30-34)	-1.619	-1.340	-1.975	-1.360	-3.030**	-2.365*	-3.855***	-3.029**	-4.347***	-3.929***	-4.131***	-4.688***
Dropout rate	-2.472	-2.588	-3.020	-2.374	-3.459**	-3.181*	-3.670***	-3.787**	-3.168***	-3.826***	-3.000***	-2.316**
Unemployment rate	0.341	0.429	0.171	0.307	-0.105	-0.054	-0.047	0.016	-1.261*	0.062	-2.347***	-2.327***
Firms (1,000 inh.)	-1.886	-2.093	-1.680	-1.727	-1.778	-1.863	-1.382	-1.865	0.169	-1.335	1.446*	0.847
Small firms (%)	0.002	0.001	0.003	0.003	0.000	0.001	-0.003	0.002	-0.006	-0.003	-0.002	-0.005
Construction firms (%)	0.010	0.012	0.004	0.008	-0.008	0.000	-0.010	-0.010	-0.010	-0.009	-0.014***	-0.009
Female mayor	-0.005	-0.001	0.004	-0.004	0.012	-0.003	0.029	0.012	0.021	0.038	-0.003	0.018
Graduated mayor	-0.052	-0.063	-0.025	-0.046	0.010	-0.025	0.029	0.023	0.034	0.013	0.038	0.037
Incumbent mayor	-0.076	-0.057	-0.074	-0.086	-0.088	-0.058	-0.091	-0.106	-0.067	-0.097	-0.025	-0.066

▸ Back

Placebo: Diff-in-Disc



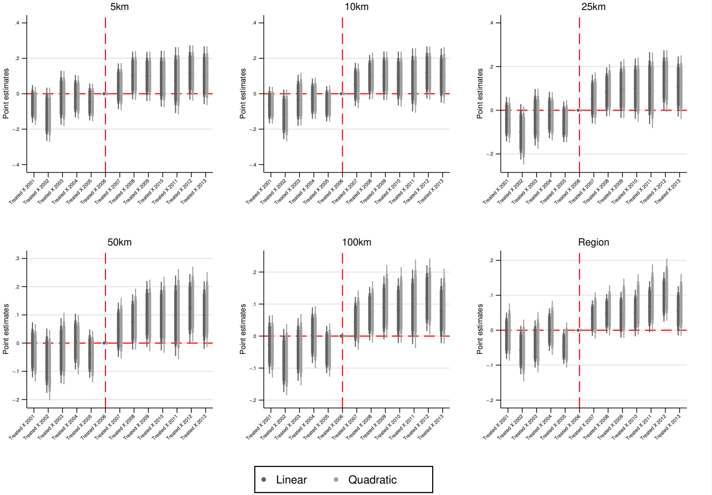
Back

Robust Coefficient: Diff-in-Disc

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.
Treat X Post 2006	0.132***	0.135***	0.141***	0.139***	0.138***	0.142***	0.115***	0.130***	0.112***	0.125***	0.083***	0.096***
	(0.041)	(0.041)	(0.040)	(0.040)	(0.037)	(0.039)	(0.033)	(0.037)	(0.027)	(0.032)	(0.017)	(0.022)
Observations	1,298	1,298	2,023	2,023	3,934	3,934	7,105	7,105	10,904	10,904	17,111	17,111
R-squared	0.706	0.708	0.700	0.702	0.724	0.724	0.784	0.784	0.823	0.823	0.806	0.806
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	2nd	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.0887	.0967	.0648	.0702	.0502	.0505	.0453	.0468	.0455	.0458	.0495	.0499
Bootstrap	.0018	.0013	.001	.0011	.0004	.0004	.0009	.001	.0001	.0002	0	0
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	.17	.17	.15	.15	.12	.12	.11	.11	.11	.11	.1	.1

Back

Robust Coefficient: Event Study



Back

Random Seeds

	(1) Inf.	(2) Inf.	(3) Inf.	(4) Inf.	(5) Inf.	(6) Inf.	(7) Inf.	(8) Inf.	(9) Inf.	(10) Inf.	(11) Inf.	(12) Inf.
Treat X Post 2006	0.102***	0.093**	0.099***	0.101***	0.089***	0.099***	0.084***	0.086***	0.065***	0.086***	0.050***	0.065***
	(0.037)	(0.038)	(0.036)	(0.037)	(0.031)	(0.035)	(0.024)	(0.031)	(0.018)	(0.025)	(0.011)	(0.016)
Observations	1,298	1,298	2,023	2,023	3,934	3,934	7,105	7,105	10,904	10,904	17,111	17,111
R-squared	0.745	0.748	0.751	0.751	0.786	0.786	0.842	0.842	0.873	0.874	0.863	0.863
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	2nd	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.0709	.0831	.0616	.0618	.0578	.059	.0619	.062	.0663	.0676	.075	.0757
Bootstrap	.0073	.0142	.0072	.0091	.0045	.0054	.0006	.0064	.0004	.0008	0	0
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	.16	.16	.14	.14	.1	.1	.1	.1	.1	.1	.09	.09

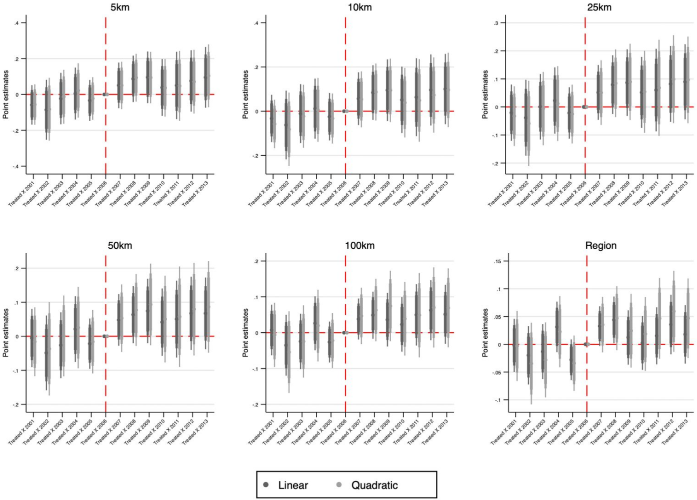
Back

Placebo: Diff-in-Disc with Sardinia

	(1)
	Inf.
Treat X Post 2006	-0.031***
	(0.006)
Observations	9,501
R-squared	0.563
Within R-squared	.0173
Bootstrap	0
Mun. FE	YES
Year FE	YES
Base value 2006	.08

Back

Removing Capital Spending from the Prediction



Back

Removing 2007-2013 from the Training

	(1) Inf.	(2) Inf.	(3) Inf.	(4) Inf.	(5) Inf.	(6) Inf.	(7) Inf.	(8) Inf.	(9) Inf.	(10) Inf.	(11) Inf.	(12) Inf.
Treat X Post 2006	0.094**	0.082**	0.097**	0.093**	0.079**	0.096**	0.060**	0.076**	0.035*	0.061**	0.002	0.020
	(0.039)	(0.040)	(0.038)	(0.039)	(0.033)	(0.037)	(0.026)	(0.033)	(0.020)	(0.027)	(0.013)	(0.017)
Observations	1,298	1,298	2,023	2,023	3,926	3,926	7,042	7,042	10,791	10,791	16,991	16,991
R-squared	0.551	0.557	0.543	0.544	0.573	0.574	0.662	0.663	0.717	0.717	0.710	0.710
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.0492	.0632	.0479	.0496	.0397	.0419	.036	.0371	.0368	.0386	.0339	.0345
Bootstrap	.0176	.0443	.0115	.0186	.0168	.011	.0225	.0219	.0815	.0274	.9027	.2449
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	.12	.12	.1	.1	.08	.08	.07	.07	.06	.06	.05	.05

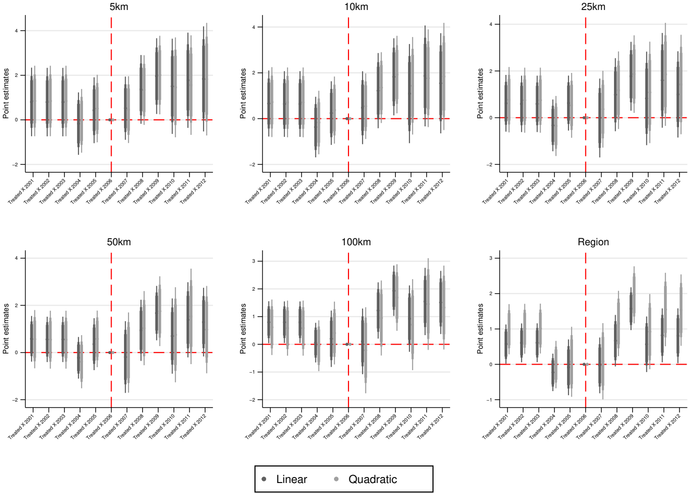
Back

Mafia-Related Crimes: Diff-in-Disc

	(1) Mafia	(2) Mafia	(3) Mafia	(4) Mafia	(5) Mafia	(6) Mafia	(7) Mafia	(8) Mafia	(9) Mafia	(10) Mafia	(11) Mafia	(12) Mafia
Treat X 2007 - 2008	0.502	0.396	0.548	0.559	0.137	0.364	0.072	0.209	0.224	0.013	0.075	0.112
	(0.463)	(0.486)	(0.520)	(0.470)	(0.550)	(0.541)	(0.487)	(0.556)	(0.374)	(0.500)	(0.243)	(0.334)
Treat X 2009 - 2012	1.327**	1.212*	1.281**	1.351**	0.950*	1.150*	0.959**	0.945*	1.040***	0.940**	0.697***	0.984***
	(0.625)	(0.665)	(0.620)	(0.627)	(0.552)	(0.617)	(0.426)	(0.552)	(0.320)	(0.443)	(0.194)	(0.270)
Observations	1,198	1,198	1,867	1,867	3,630	3,630	6,567	6,567	10,072	10,072	15,795	15,795
R-squared	0.476	0.479	0.459	0.461	0.425	0.426	0.425	0.426	0.448	0.449	0.460	0.461
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	2nd	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.2484	.252	.2368	.2398	.2147	.216	.2184	.22	.2411	.2418	.2506	.2519
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	2.14	2.14	2.14	2.14	2.21	2.21	2.33	2.33	2.09	2.09	1.84	1.84

Back

Mafia-Related Crimes: Event Study



Back

Mafia-Related Crimes: Diff-in-Disc - 2 Post Periods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Mafia	Mafia	Mafia	Mafia	Mafia	Mafia	Mafia	Mafia	Mafia	Mafia	Mafia	Mafia
Treat X 2007 - 2008	0.502	0.396	0.548	0.559	0.137	0.364	0.072	0.209	0.224	0.013	0.075	0.112
	(0.463)	(0.486)	(0.520)	(0.470)	(0.550)	(0.541)	(0.487)	(0.556)	(0.374)	(0.500)	(0.243)	(0.334)
Treat X 2009 - 2012	1.327**	1.212*	1.281**	1.351**	0.950*	1.150*	0.959**	0.945*	1.040***	0.940**	0.697***	0.984***
	(0.625)	(0.665)	(0.620)	(0.627)	(0.552)	(0.617)	(0.426)	(0.552)	(0.320)	(0.443)	(0.194)	(0.270)
Observations	1,198	1,198	1,867	1,867	3,630	3,630	6,567	6,567	10,072	10,072	15,795	15,795
R-squared	0.476	0.479	0.459	0.461	0.425	0.426	0.425	0.426	0.448	0.449	0.460	0.461
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	2nd	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.2484	.252	.2368	.2398	.2147	.216	.2184	.22	.2411	.2418	.2506	.2519
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	2.14	2.14	2.14	2.14	2.21	2.21	2.33	2.33	2.09	2.09	1.84	1.84

Back

Only Decentralization: Lazio Vs Campania

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.	Inf.
Treat X Post 2006	0.124	0.094	0.127	0.128	0.074	0.118	0.043	0.057	-0.002	0.051	0.018	-0.026
	(0.109)	(0.097)	(0.110)	(0.109)	(0.087)	(0.107)	(0.055)	(0.085)	(0.032)	(0.054)	(0.018)	(0.032)
Observations	273	273	491	491	1,137	1,137	2,793	2,793	7,323	7,323	12,002	12,002
R-squared	0.786	0.798	0.800	0.801	0.839	0.841	0.872	0.872	0.862	0.862	0.835	0.836
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	2nd	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.0981	.1462	.0782	.0827	.0458	.0546	.0462	.0466	.0409	.0433	.0434	.0449
Bootstrap	.256	.3364	.2692	.242	.4418	.3097	.4654	.5645	.9368	.3714	.3138	.4222
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	.06	.06	.08	.08	.07	.07	.09	.09	.1	.1	.09	.09

Back

Decentralization + Increase funding: Puglia Vs Molise

	(1) Inf.	(2) Inf.	(3) Inf.	(4) Inf.	(5) Inf.	(6) Inf.	(7) Inf.	(8) Inf.	(9) Inf.	(10) Inf.	(11) Inf.	(12) Inf.
Treat X Post 2006	0.184***	0.192***	0.198***	0.188***	0.187***	0.212***	0.142***	0.181***	0.087**	0.146***	0.093***	0.119***
	(0.057)	(0.059)	(0.053)	(0.056)	(0.047)	(0.052)	(0.041)	(0.047)	(0.036)	(0.043)	(0.029)	(0.037)
Observations	324	324	506	506	1,078	1,078	1,920	1,920	2,658	2,658	5,109	5,109
R-squared	0.465	0.469	0.477	0.478	0.496	0.498	0.505	0.510	0.537	0.541	0.538	0.539
Poly.	1st	2nd	1st	2nd	1st	2nd	1st	2nd	2nd	2nd	1st	2nd
Specification	5Km	5Km	10Km	10Km	25Km	25Km	50Km	50Km	100Km	100Km	Regions	Regions
Within R-squared	.137	.1426	.1242	.1268	.0828	.0873	.042	.0518	.0362	.0437	.0794	.0818
Bootstrap	.0033	.003	.0005	.0037	.0001	0	.0014	.0006	.0232	.0023	.0009	.0023
Mun. FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Base value 2006	.32	.32	.26	.26	.17	.17	.13	.13	.12	.12	.12	.12

Back

Bibliography I

- Acconcia, A., Corsetti, G., & Simonelli, S. (2014).Mafia and Public Spending: Evidence on the Fiscal Multiplier from a Quasi-experiment. *American Economic Review*, *104*(7), 2185–2209. https://doi.org/10.1257/aer.104.7.2185
- Acemoglu, D., Robinson, J. A., & Santos, R. J. (2013).The monopoly of violence: Evidence from colombia. *Journal of the European Economic Association*, *11*(suppl_1), 5–44.
- Alesina, A., & Perotti, R. (2002).The european union: A politically incorrect view. *Journal of Economic Perspectives*, *18*(4), 27–48.
- Alesina, A., Piccolo, S., & Pinotti, P. (2019).Organized Crime, Violence, and Politics. *The Review of Economic Studies*, *86*(2), 457–499. https://doi.org/10.1093/restud/rdy036
- Ash, E., Galletta, S., & Giommoni, T. (2020). A Machine Learning Approach to Analyzing Corruption in Local Public Finances. https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/414608/1/CLE_WP_2020_06.pdf

Bibliography II

Athey, S. (2018). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda* (pp. 507–547). University of Chicago Press.

Battiston, G., Daniele, G., Le Moglie, M., & Pinotti, P. (2022).Fueling organized crime: The mexican war on drugs and oil thefts.

Becker, S. O., Egger, P. H., & Von Ehrlich, M. (2010a).Going nuts: The effect of eu structural funds on regional performance. *Journal of Public Economics*, *94*(9-10), 578–590.

Becker, S. O., Egger, P. H., & Von Ehrlich, M. (2010b).Going nuts: The effect of eu structural funds on regional performance. *Journal of Public Economics*, *94*(9-10), 578–590.

Butts, K. (2023).Geographic difference-in-discontinuities [Publisher: Routledge _eprint: https://doi.org/10.1080/13504851.2021.2005236]. *Applied Economics Letters*, *30*(5), 615–619. https://doi.org/10.1080/13504851.2021.2005236

Castillo, J. C., & Kronick, D. (2020).The logic of violence in drug war. *American Political Science Review*, *114*(3), 874–887.

Bibliography III

- Daniele, G., & Dipoppa, G. (2017).Mafia, elections and violence against politicians. *Journal of Public Economics*, *154*, 10–33. https://doi.org/10.1016/j.jpubeco.2017.08.004
- Daniele, G., & Dipoppa, G. (2023).Fighting Organized Crime by Targeting their Revenue: Screening, Mafias, and Public Funds. *The Journal of Law, Economics, and Organization*, *39*(3), 722–746. https://doi.org/10.1093/jleo/ewac002
- Daniele, G., & Geys, B. (2015).Organised Crime, Institutions and Political Quality: Empirical Evidence from Italian Municipalities [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12237]. *The Economic Journal*, *125*(586), F233–F255. https://doi.org/10.1111/ecoj.12237
- Dell, M. (2015).Trafficking networks and the mexican drug war. *American Economic Review*, *105*(6), 1738–1779.
- Dube, O., & Vargas, J. F. (2013).Commodity price shocks and civil conflict: Evidence from colombia. *Review of Economic studies*, *80*(4), 1384–1421.

Bibliography IV

- Fenizia, A. (2018, January). Breaking the Ties between the Mafia and the State: Evidence from Italian Municipalities. https://doi.org/10.2139/ssrn.3105798
- Galletta, S. (2017).Law enforcement, municipal budgets and spillover effects: Evidence from a quasi-experiment in Italy. *Journal of Urban Economics*, *101*, 90–105. https://doi.org/10.1016/j.jue.2017.06.005
- Glaeser, E. L., Hillis, A., Kominers, S. D., & Luca, M. (2016).Crowdsourcing city government: Using tournaments to improve inspection accuracy. *American Economic Review*, *106*(5), 114–118.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018).Human decisions and machine predictions. *The quarterly journal of economics*, *133*(1), 237–293.
- Lessing, B. (2017). *Making peace in drug wars: Crackdowns and cartels in latin america*. Cambridge University Press.

Bibliography V

- Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., & Brantingham, P. J. (2015).Randomized controlled field trials of predictive policing. *Journal of the American statistical association*, *110*(512), 1399–1411.
- Organized Crime Index, O. C. I. (2023). *Organized crime index 2023* (tech. rep.). Global Initiative against Transnational Organized Crime.
- Pinotti, P. (2015a).The causes and consequences of organised crime: Preliminary evidence across countries [Publisher: [Royal Economic Society, Wiley]]. *The Economic Journal*, *125*(586), F158–F174. Retrieved March 13, 2022, from https://www.jstor.org/stable/24737564
- Pinotti, P. (2015b).The Economic Costs of Organised Crime: Evidence from Southern Italy [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12235]. *The Economic Journal*, *125*(586), F203–F232. https://doi.org/10.1111/ecoj.12235
- Sviatschi, M. M. (2022).Making a narco: Childhood exposure to illegal labor markets and criminal life paths. *Econometrica*, *90*(4), 1835–1878.

Bibliography VI

United Nations Office on Drugs and Crime. (2016).Legislative guide for the implementation of the united nations conventions against transnational organized crime.