DISCUSSION PAPER SERIES

DP19322

MAFIA, POLITICS AND MACHINE PREDICTIONS

Gian Maria Campedelli, Gianmarco Daniele and Marco Le Moglie

POLITICAL ECONOMY AND PUBLIC ECONOMICS

CEPR

ISSN 0265-8003

MAFIA, POLITICS AND MACHINE PREDICTIONS

Gian Maria Campedelli, Gianmarco Daniele and Marco Le Moglie

Discussion Paper DP19322
Published 02 August 2024
Submitted 22 July 2024

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Political Economy
- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Gian Maria Campedelli, Gianmarco Daniele and Marco Le Moglie

MAFIA, POLITICS AND MACHINE PREDICTIONS

Abstract

Detection is one of the main challenges in the fight against organized crime. We show that machine learning can be used to predict mafias infiltration in Italian local governments, as measured by the dismissal of city councils infiltrated by organized crime. The model successfully predicts up to 96% of out-of-sample municipalities previously identified as infiltrated by mafias, up to two years earlier, making this index a valuable tool for identifying municipalities at risk of infiltration well in advance. Furthermore, we can identify "high-risk" local governments that may be infiltrated by organized crime but have not been detected by the state, thereby improving the efficacy of detection. We then apply this new time-varying measure of organized crime to investigate the underlying causes of this type of rent-seeking. As criminals infiltrate politics to capture public resources, we study how a positive shock in public spending (European Union transfers), affects this phenomenon. Employing a geographic Difference-in-Discontinuities design, we find a substantial and lasting increase in the predicted risk of mafia infiltration (up to 14 p.p.), emphasizing the unintended effects of delivering aid where criminal organizations can appropriate public funds.

JEL Classification: D72, K4

Keywords: Crime, Corruption, Politics, Elections

Gian Maria Campedelli - gcampedelli@fbk.eu
Fondazione Bruno Kessler

Gianmarco Daniele - daniele.gianmarco@gmail.com
University of Milan and CEPR

Marco Le Moglie - marco.lemoglie@unicatt.it
Catholic University of the Sacred Heart

Acknowledgements
This research has been funded by PRIN 2022ZNCLYL "Fighting Organized Crime" granted by the Italian Ministry of University and Research. We are thankful for comments received from Alessandra Fenizia, Sergio Galletta, Paolo Pinotti, Maria Petrova, Federico Masera, Gabriele Gratton, Adam Soliman, Piero Stanig, Massimo Pulejo, Gemma Dipoppa, Raffaele Lagravinese, Alberto Aziani, Davide Cipullo, Giovanna Marcolongo, and Zubin Jelveh. We also benefited from suggestions and feedback received from participants at seminars held at Bocconi University – CLEAN Unit, the ETH Center for Law and Economics, Transcrime, Università Cattolica del Sacro Cuore, Fondazione Bruno Kessler, University of Bergen, UNSW Business School, the 2nd Resilient Democracy Lab Workshop in Sydney, Univeridad Carlos III de Madrid, Institut d'Economia de Barcelona, and the 8th Workshop on the Economics of Organized Crime at Collegio Carlo Alberto.

Mafia, Politics and 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]Catholic University of the Sacred Heart

July 22, 2024

Abstract

Detection is one of the main challenges in the fight against organized crime. We show that machine learning can be used to predict mafias infiltration in Italian local governments, as measured by the dismissal of city councils infiltrated by organized crime. The model successfully predicts up to 96% of out-of-sample municipalities previously identified as infiltrated by mafias, up to two years earlier, making this index a valuable tool for identifying municipalities at risk of infiltration well in advance. Furthermore, we can identify "high-risk" local governments that may be infiltrated by organized crime but have not been detected by the state, thereby improving the efficacy of detection. We then apply this new time-varying measure of organized crime to investigate the underlying causes of this type of rent-seeking. As criminals infiltrate politics to capture public resources, we study how a positive shock in public spending (European Union transfers), affects this phenomenon. Employing a geographic Difference-in-Discontinuities design, we find a substantial and lasting increase in the predicted risk of mafia infiltration (up to 14 p.p.), emphasizing the unintended effects of delivering aid where criminal organizations can appropriate public funds.

*This research has been funded by PRIN 2022ZNCLYL "Fighting Organized Crime" granted by the Italian Ministry of University and Research. We are thankful for comments received from Alessandra Fenizia, Sergio Galletta, Paolo Pinotti, Maria Petrova, Federico Masera, Gabriele Gratton, Adam Soliman, Piero Stanig, Massimo Pulejo, Gemma Dipoppa, Raffaele Lagravinese, Alberto Aziani, Davide Cipullo, Giovanna Marcolongo, and Zubin Jelveh. We also benefited from suggestions and feedback received from participants at seminars held at Bocconi University – CLEAN Unit, the ETH Center for Law and Economics, Transcrime, Università Cattolica del Sacro Cuore, Fondazione Bruno Kessler, University of Bergen, UNSW Business School, the 2nd Resilient Democracy Lab Workshop in Sydney, Univeridad Carlos III de Madrid, Institut d'Economia de Barcelona, and the 8th Workshop on the Economics of Organized Crime at Collegio Carlo Alberto.

1 Introduction

Criminal organizations pose a serious threat to societies worldwide, as their presence reduces economic growth (Pinotti, 2015) and human capital accumulation (Sviatschi, 2022a), while increasing political violence (Daniele and Dipoppa, 2017) and disrupting legal businesses (Mirenda et al., 2022; Le Moglie and Sorrenti, 2022). This pervasive influence often extends to politics, characterized by the strategic goal of manipulating elections, capturing public resources, and opposing crackdown efforts (Acemoglu et al., 2013; Alesina et al., 2019).[1]

While countering criminal organizations represents a priority for governments around the world, curbing their power remains a challenging goal. A key obstacle in tackling organized crime is the difficulty in uncovering its activities and presence. Detection is generally costly and imperfect, and this is especially true for organized crime, which involves complex and underreported crimes like corruption or money laundering (Cressey, 1967; United Nations Office on Drugs and Crime, 2016; Le Moglie and Sorrenti, 2022).[2] In this paper, we address this issue by showing how machine learning can be used to improve detection of organized crime, and in turn, the state capability to fight criminal organizations.

We study this phenomenon in Italy, by focusing on one the most assertive policies against organized crime: the dismissal of city councils. This disruptive policy tool allows the central government to regain control and restore legitimacy in areas where the mafia has significant influence over local politics. The central government can dismiss the entire municipal government and appoint a temporary technocratic government, if there is evidence of ties between criminal organizations and local politicians.

We leverage machine learning to predict which municipal governments will be detected as infiltrated by organized crime. This serves two purposes: first, we can accurately anticipate state detection of organized crime infiltration in local governments, increasing the efficiency of detection. Second, we can highlight "high risk" local governments possibly infiltrated by

[1] According to a recent report by the Global Organised Crime Index (2021), an alarming 80% of the world's population resides in countries where the risk posed by organized crime is notably high.

[2] As detection is limited, it is therefore difficult to measure organized crime activities. Generally, their presence is measured *ex post* when police interventions crack down on criminal activities, such as seizing drugs and arrests, or when there are visible signs of their presence, i.e. violence, often measured by homicides or attacks (Dell, 2015; Sviatschi, 2022b; Limodio, 2022).

organized crime that have not been detected by the state, increasing the efficacy of detection.

We predict (detected) mafia infiltration exploiting information extracted from local elections and budget accounts spanning the years 2001 to 2020, as mafia interest in local politics involves the misappropriation of public resources and influence on local elections (Daniele and Geys, 2015; Di Cataldo and Mastrorocco, 2022; Fenizia and Saggio, 2023). Overall, we evaluate 200 combinations of model architectures sampling from more than 1,500 candidates, comparing nine different algorithms.[3] XGBoost models emerge as the most effective.

The best XGBoost model successfully predicts 96% of out-of-sample local governments detected as infiltrated by the Italian state.[4] The prediction is particularly meaningful when focusing on the most at risk local governments, according to this synthetic indicator.

To exemplify, when focusing on local governments in the top 5% of the predicted risk in a given year, we predict around 89% of all infiltrated local governments, with a precision rate of approximately 15% (i.e. one truly infiltrated local government for every five false positives). In comparison, a random selection of 5% of all local governments would identify only about 4% of infiltrated local governments or 18% with a draw from Southern regions, where this phenomenon is more widespread. The predictive performance remains comparable when using the estimated risk to predict infiltration one year and two years ahead, where we respectively identify 97.4% and 95.9% of all infiltrated local governments: This implies that we are able to predict infiltration up to two years before it is detected by the Italian state.

Importantly, false positives are not necessarily incorrect, as they may represent infiltrated local governments that have not been detected by the Italian state. As our prediction is based on the state detection capability, we argue that it is highly unlikely for authorities to identify all infiltrated local governments, given the widespread use of corruption involved in these activities. This implies that local governments identified as infiltrated by our index but not by the Italian state may, to some extent, represent a subset of infiltrated councils that have evaded

[3]The nine tested algorithms are: Logistic Regression, Lasso Regression, Ridge Regression, Elastic Net, Neural Networks, Decision Trees, Random Forests, Gradient Boosting Machines, and XGBoost.

[4]This is based on the conventional 0.5 probabilistic threshold.

detection.[5] To provide some indirect evidence, we design a comparison across three groups of local governments: mafia (detected) infiltrated ones; a subset of these false positives, where we suspect a very high risk of infiltration (*suspected*); all other local governments. Controlling for region-year and municipality fixed effects, we find that infiltrated local governments are more likely to experience a mafia murder or a violent attack against a local politician. More interestingly, we observe similar patterns among the *suspected* governments, supporting the idea that these municipalities are indeed at high risk of infiltration.

Overall, our computational approach provides a novel tool for anti-organized crime policies. Similarly to what happens with corruption, these investigations are complex and rely on limited information. We suggest that investigators can enrich their set of information by considering machine predictions. While recent papers show the relevance of machine learning in relation to violent crimes (Kleinberg et al., 2018; Heller et al., 2022; Campedelli, 2022a), we believe this approach can be useful also for more complex crimes like organized crime.

An additional benefit of this approach is that our index can be used as a local, time-varying measure of organized crime's influence on local politics. This measure can then be employed to evaluate how external shocks impact this phenomenon. In the paper, we apply this reasoning to investigate the causes of organized crime infiltration: as criminals infiltrate politics to capture public resources, we examine how a positive shock in public spending affects criminal influence on politics.

We focus on the impact of European funds aimed at fostering economic growth in disadvantaged areas across the European Union. This redistributive policy is the most important source of public investments for many local governments in the European Union (Becker et al., 2010). On the one hand, EU funds could reduce organized crime influence on politics as they might foster economic growth, thereby decreasing local dependence on mafia-provided services and their sway over politics. On the other hand, criminal organizations might infiltrate local governments that are receiving substantial amounts of EU funds, seeking to divert these funds for their gain.

[5] Another implication of the limited detection is that we can successfully predict mafia infiltration when local governments mimic the patterns observed in detected municipalities. Conversely, we cannot predict infiltration in undetected local governments that present patterns completely unrelated to the detected ones.

The identification is based on a significant surge in EU financial aid to Southern Italian municipalities, which coincided with a reduction in fiscal oversight during 2007-2013, involving a total of EUR 56 billion. Using a geographic Differences-in-Discontinuities approach, we contrast municipalities near regional borders with/without these funds, comparing periods before and after 2007. Our analysis employs various proximity thresholds, ranging from 5km to 100km from the border.

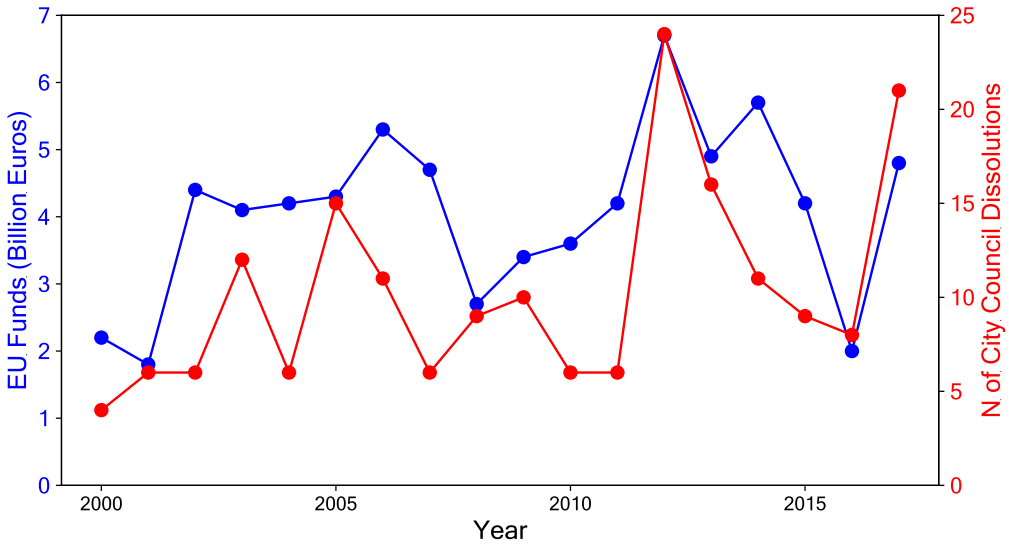
Ideally, we could explore this question by using local governments that were actually dismissed due to mafia infiltration as the dependent variable. However, although Figure 1 indicates a significant correlation between dismissed municipalities and EU funds disbursement (0.63, p-val=0.005), the very small number of detected local governments in the period and area of interest hinders this approach. Conversely, our predicted index, which is highly granular and time-sensitive, represents a viable option to investigate this question. In line with the correlational evidence, we identify a pronounced and lasting increase in the likelihood of mafia infiltration after funds are distributed to entitled municipalities, with an increase of 11 to 14 percentage points in our predictive indicator.[6]

Results are confirmed when using an overall measure of mafia presence, i.e. mafia crimes in a municipality. Notably, however, the impact of our index is observed immediately following the disbursement of funds, whereas mafia charges – linked to the state's capacity to identify mafia activities – show an increase after a delay of one to two years. This further corroborates that our index offers a more timely detection of mafia presence, significantly ahead of traditional detection methods.

This paper contributes to the literature on organized crime highlighted in the opening paragraph. While there is abundant evidence highlighting the negative effects of organized crime, the literature on successful policies to counteract it is critically sparse. Extant examples, primarily concentrated on Latin America, show controversial results associated with crackdowns (Dell, 2015; Battiston et al., 2022), deportation policies (Sviatschi, 2022b), and heightened policing (Blattman et al., 2021), often resulting in an unintended increase in gang influence. Instead, evidence suggests that the city council dismissal could represent an interesting excep-

[6]In a set of robustness tests, we exclude the direct influence of European funds on our index (i.e. variations in public spending due to the receipt of EU funds that might mechanically affect our index).

FIGURE 1. EU Funds and City Council Dismissals



Note: This figure visualizes the trends (from 2000 to 2017) in terms of yearly number of city council dismissals (red line) and EU funds assigned to Italy, in Billion Euros (blue line). The Pearson correlation between the two is 0.63 (p-val=0.005).

tion: Previous works have shown the indirect positive effects of this policy on political selection (Daniele and Geys, 2015; Baraldi et al., 2022), economic growth (Fenizia and Saggio, 2023) and investment decisions in neighboring areas (Galletta, 2017). In this paper, we show how machine learning can be used to improve the enforcement of this policy both in terms of efficiency and efficacy, providing a favorable benefit-cost ratio (as discussed later).

Second, we contribute to the literature on the determinants of organized crime (Acemoglu et al., 2020; Buonanno et al., 2015). By using this index to investigate the impact of EU funds on mafia infiltration, we highlight the potential downsides of policies aimed at stimulating economic growth in disadvantaged regions: A combination of transfers, along with a reduction in monitoring, can unintentionally lead to the expansion of mafia influence. This is in line with previous evidence on criminals' ability to infiltrate and capture subsidies to firms and public procurements (Barone and Narciso, 2015; Marcolongo, 2023; Daniele and Dipoppa, 2023). In fact, our results are in line with a "rapacity effect", wherein the inflow of EU funds change the incentives to steal resources from local governments. A parallel can be drawn to Colombia where Dube and Vargas (2013) show that higher oil prices fuelled conflict among criminals

to appropriate contestable resources. In our scenario, the mechanism differs, with corruption directly implicating local politicians instead of straightforward theft.[7]

Finally, our contribution extends to the growing literature on the application of machine learning to crime-related problems (Campedelli, 2022b). Recent scholarship showcases the relevance of machine learning in crime prevention, such as aiding bail decisions (Kleinberg et al., 2018), preventing shooting victimization (Heller et al., 2022), studying homicide clearance (Campedelli, 2022a) and identifying politically connected firms (Titl et al., 2024). This paper is the first to apply machine learning to predict organized crime, a complex and under-reported crime often intertwined with corruption. This connects our work to recent studies that use machine learning to predict corruption among Brazilian local governments (Ash et al., 2020; Colonnelli et al., 2020).[8]

In the next section, we describe the Institutional Background. The machine learning approach is detailed in Section 3. In Section 4, we study whether the arrival of EU funds impacts mafia infiltration into local politics. We conclude in Section 5.

2 Institutional Background

2.1 Local Governance in Italy

In the administrative structure of Italy, municipal governments represent the lowest tier. As of 2023, Italy counts 7,901 municipalities, each presided over by a mayor, an executive committee, and a municipal council. Municipal elections take place every five years, with mayors directly elected for up to two consecutive terms (except in municipalities with fewer than 3,000

[7] Becker et al. (2010) show that EU funds fostered economic growth in many European regions, while having a limited impact in Southern Italy. Our findings might rationalize this puzzle, suggesting that such policies can have unintended effects in areas with strong organized crime presence.

[8] We methodologically contribute to this field by underscoring the importance of algorithmic comparison and extensive hyperparameter grid search to contrast the risks of weak predictive pipelines. We also show that artificial upsampling techniques can reduce issues from class imbalance, and highlight the benefits deriving from coupling algorithmic prediction with explanation (Hofman et al., 2021).

inhabitants, where – since 2014 – they can serve up to three terms).[9] Municipal councils play a pivotal role in governing various policy domains such as local transportation, local policing, social welfare, waste management, and local infrastructures (e.g. road maintenance). The municipal government has also some limited power in setting local taxes, e.g. regarding waste collection.

2.2 City Council Dismissals for Mafia Infiltration

The policy of city council dismissal, introduced in response to the escalating influence of the mafia on local governments in the 1980s, was enacted by the Italian parliament in 1991 (D.L. 31/05/1991 n. 164). This extraordinary preventive measure applies when there is a real danger that the activity of a municipality is bent on the interests of mafia clans. The legislation empowers the central government to replace the mayor, executive committee, and city council with external commissioners. These commissioners are experienced career civil servants, who assume full executive and legislative powers, governing the municipality for a duration of 12 to 24 months until new elections are held.

The dismissal process can be split in three steps:

1. **Preliminary Investigations**: Provincial branches of the national police forces may discover potential connections between criminals and local politicians during their ordinary activities, either accidentally, during unrelated investigations (e.g., intercepting phone calls of individuals already under scrutiny) or due to specific events (e.g., a mafia-style attack against a politician). Our contribution focuses on this stage. We propose a tool that investigators can use to enrich their set of information on local governments at risk.

2. **Official Investigation**: If preliminary investigations reveal a credible threat, the police requests an official investigation from the Ministry of Interior. The dismissal occurs after a six-month investigation conducted by a special commission appointed by the Ministry

[9] Municipalities with more than 15,000 inhabitants may have a runoff round if no mayoral candidate reaches the absolute majority of votes (50%+1) in the first round. For municipalities with less than 15,000 inhabitants, the electoral law mandates a single electoral round where the mayoral candidate reaching the relative majority of votes is automatically elected.

of Interior.[10] The law prescribes a rigorous procedure for the dismissal process, from the emergence of evidence to the final decision, to minimize arbitrariness and delays.

3. **Post-Dismissal Review**: The appointed commissioners are responsible for reviewing the policies enacted by the dismissed municipal council. This includes halting new projects, conducting thorough financial reviews, and revoking contracts obtained unlawfully or through connections with organized crime.

Since its enactment in 1991, the law has been invoked 378 times. In some municipalities, it has been invoked multiple times, underscoring that in environments where organized crime has significant influence, this policy alone is insufficient to eliminate mafia control over politics. In a few instances, dismissals have been revoked following subsequent investigations, which revealed the absence of connections between criminals and local politicians.

Figure S1 in the Supplementary Materials shows the distribution of city councils dismissed for mafia infiltration. Most dismissals occurred in the three regions traditionally associated with mafia presence: Calabria, Campania, and Sicily. However, in recent years, the law has also been applied in some municipalities in Northern Italy.

Notably, the dismissals are not triggered by poor municipal financial performance or inefficiencies in public procurement but are instead responsive to criminal activities such as extortion, drug and arms trafficking, money laundering, vote buying, and collusion in public procurement (Fenizia and Saggio, 2023).

3 Methods

3.1 Data Selection

The dataset upon which this work relies uniquely combines administrative data publicly available from the *Italian Ministry of Interior* from 2000 to 2020. We focus on electoral and public

[10] Since 2017, 94% of investigations ended in dismissal. Only in two cases, Avola and San Giovanni di Gerace (see https://www.interno.gov.it/it/notizie/avola-insediata-commissione-indagine-antimafia), did the investigation find no evidence of mafia infiltration.

spending variables for two main reasons. First, criminal organizations aim to influence policy-making by corrupting politicians or manipulating elections. They gain by directing procurements to firms they control, influencing hiring decisions, and gaining privileged access to local services, e.g. building permits (Fenizia and Saggio, 2023; Di Cataldo and Mastrorocco, 2022). Second, these variables are publicly available and easily accessible, enabling straightforward updates to this index in the future, a very relevant aspect for policy applications. Conversely, we refrain from using potentially biased data that are a direct byproduct of detection (e.g. anti-mafia police operations) or data that are only available only for recent years (e.g. detailed procurement data).

Politics. First, we collect political and electoral data mapping electoral features and outcomes. From this dataset, we identified seven crucial variables extensively studied in the political corruption literature: (1) share of votes for the winning mayoral candidate, (2) female mayor, (3) number of competitors, (4) mayors' incumbency and (5) education (a dummy = 1 if the mayor holds a college degree), (6) a dummy for whether the mayor was born in the same municipality of election, and (7) municipal population.[11] Ideology variables refer to the prevalent political background of the winning list. We have created six binary variables covering the entire political spectrum: (1) far right, (2) right, (3) civic list, (4) left, (5) far left, (6) M5s.[12]

Public Spending. Second, we collect municipal public spending data. Municipal public spending spans across twenty-two distinct categories,[13] and each one has two types of spending: current expenditures (*conto corrente*) and capital expenditures (*conto capitale).* Current expenses pertain to current or operational expenditures, representing the day-to-day expenses

[11]Municipal population was transformed using the min-max scaler in Scikit-learn: We re-scale it to avoid issues with variables having large variance and excessively heterogeneous scales in the training phase of the algorithms.

[12]Far right includes parties such as the *Alleanza Nazionale*, a post-fascist nationalist party founded in 1994 (and dissolved in 2009). Right includes parties such as *Forza Italia*, a center-right party funded by Silvio Berlusconi in 1994. Civic lists include all those local parties that are not associated with national parties. Left lists include, for instance, *Partito Democratico*, the center-left main Italian party. Far-left parties include several minor extreme-left parties, e.g. *the Italian Communist Party.* Finally, M5S refers to lists being associated with the *Movimento 5 Stelle*, a populist party refusing the typical left-right ideology scale, founded in 2009.

[13]Agriculture, Debt Interests, Elderly Assistance, Culture, Debt, Development, Education, Energy, Environment, Provisional Funds, Administration, Health, Justice, Employment, Debts with other Public Institutions, Social Welfare, Local Police, Sport, Urban Planning, Tourism, Transportation.

necessary to maintain and run government services and operations. They entail operational costs such as salaries, utilities, rent, maintenance, and other routine expenditures. Capital expenses instead are related to investments. These are typically long-term investments that contribute to the development or acquisition of assets or infrastructure.

Building upon these 44 raw variables, we have created features that capture various aspects of how each spending item influences the overall expenditure of a municipality in a particular year both at time t and at time t-1.

First, for each item, we have computed the share of current (or capital) expenses over the total amount of current (or capital) expenses in the same year.[14] The rationale is to evaluate how much each item contributes to the overall amount of respectively current or capital expenditures. Then, for each item we computed the amount of current expenses over the total expenses for that specific item.[15] The goal is to discern if an item is primarily driven by operational costs, investments, or a combination of both. Another set of features focuses on the total amount spent per 10,000 inhabitants for each spending item.[16] Additionally, we have calculated two overall variables: total current expenses per capita and total capital expenses per capita. These, along with the rate variables, have been rescaled using min-max to prevent excessive variance.

Target variable: Mafia infiltration. Finally, we collect data on city council dismissals due to mafia infiltration from 1991 to 2023. Our framework centers on predicting the detection of mafia infiltration in a specific municipality-year. We formulated our binary target variable, tailored for a binary classification task, by assigning a value of 1 to the year a city council was dismissed due to infiltration, and all preceding years in the same electoral term. This decision rests on the assumption that infiltration regards the whole electoral term up to the point of the dismissal decree. For instance, if a term theoretically spans 2007 to 2012 and the council faces dismissal in 2011 due to mafia infiltration, we coded infiltration from 2007 through 2011.

[14] For instance, for current expenses: Share Curr. $(x_i)_{m,t} = \frac{\text{Curr}(x_i)_{m,t}}{\sum_{j=1}^{N} \text{Curr}(x_j)_{m,t}}$, where i indicates the specific feature, m the municipality, t the year, and j from 1 to N the entire list of public spending variables for municipality m.

[15] Computed as: Share Curr. Global $(x_i)_{m,t} = \frac{\text{Curr}(x_i)_{m,t}}{(\text{Curr}(x_i)_{m,t} + \text{Cap}(x_i)_{m,t})}$

[16] Computed as $\text{Rate}(x_i)_{m,t} = \left(\frac{\text{Curr}(x_i)_{m,t} + \text{Cap}(x_i)_{m,t}}{\text{pop}_{m,t}} \right) \times 10k$

3.2 Predictive Approach

A Highly Unbalanced Problem. Predicting mafia infiltration in local governments represents a challenging task due to the pronounced class imbalance between infiltrated and non-infiltrated observations. The original training set exhibits a mere 0.83% share of infiltrated city councils, with 855 out of 105,628 observations labeled as such. For this reason, we deployed different artificial upsampling techniques to counter the problem of class imbalance. Specifically, we have adopted three different approaches for generating synthetic observations in the training phase. These are Synthetic Minority Over-sampling Technique (SMOTE) (Chawla et al., 2002), Adaptive Synthetic sampling (ADASYN) (Haibo He et al., 2008), and SMOTE combined with Tomek links (SMOTE+Tomek) (Batista et al., 2004).[17]

Evaluating the Model. There are several metrics to evaluate the performance of an algorithm in a machine learning classification task, each with its unique rationale. In the context of our primary task, we predominantly focus on recall, which is calculated as follows:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{1}$$

Recall identifies the proportion of positive instances correctly identified by the model, i.e. those indicating infiltrated local governments.

Our emphasis on recall stems from practical considerations. Firstly, we aim to maximize the predictive performance of our task by reaching as much as possible the same detection ability demonstrated by the Italian state. We acknowledge that privileging recall, especially in highly unbalanced problems, might lead to an undesired amount of false positives. However, we argue that the costs of failing to predict an actually infiltrated local government outweigh the costs of predicting as infiltrated a local government that was not detected as infiltrated. Nevertheless, to mitigate the rate of excessive false positives, our strategy involves concentrating on subsets of

[17]SMOTE generates synthetic observations for the minority class by interpolating between existing ones. ADASYN refines this by emphasizing sparser areas for a more detailed oversampling. SMOTE combined with Tomek links (i.e. pairs of observations, one from each class, that are close to each other in the feature space) tackles both oversampling and undersampling, removing noisy majority class observations identified by Tomek links. Additional details are in Section A.2 in the Supplementary Materials.

high-risk local governments (we elaborate on the trade-off recall-precision in the next section). This approach aims to strike a balance between the emphasis on recall and the importance of Precision. Precision is calculated as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{2}$$

It measures the accuracy of positive predictions by assessing the ratio of correctly predicted positive instances (True Positives) to the total number of instances predicted as positive (sum of True Positives and False Positives). In other words, the rationale behind precision is to provide insights into the model's ability to avoid making false positive predictions. The precision-recall trade-off is a common consideration in most classification applications in machine learning, where optimizing for one metric may impact the other, and finding the right balance depends on the specific context.

The second motivation behind our focus on recall is that infiltrated local governments are only so as a result of state detection efforts. Hence, it is plausible that additional municipalities remain infiltrated but undetected due to resource constraints on policing or the absence of visible, enforceable behaviors like mafia-related crimes. Indeed, given the likely presence of infiltrated local governments that go undetected by authorities, predicted false positives might have a relevant policy value. Offering a set of local governments at risk larger than the set of municipalities detected as infiltrated would provide investigators with a sample of local governments that appear to be characterized by political and public spending patterns that are highly similar to those that were detected as infiltrated, facilitating targeted investigations.

In simpler terms, our goal is to detect infiltrated local governments as effectively as the Italian state, without an excessive rate of false positive. Additionally, the imperfect detection implies that some of the predicted false positives could actually be infiltrated local governments that have not been identified by the Italian state.

To complement recall, in the Supplementary Materials we provide further details on the

predictive performance via the Roc-Auc score.[18]

Algorithmic Comparison. In order to optimize the predictive accuracy of our machine learning approach, we have set up a framework comprehensively comparing nine different algorithmic techniques. The tested algorithms are: Logistic Regression, Lasso Regression, Ridge Regression, Elastic Net, Neural Networks, Decision Trees, Random Forests, Gradient Boosting Machines, and XGBoost. The algorithms cover a range of different approaches to predictive tasks.[19]

Building on this list of algorithms, we have performed randomized grid search across a range of relevant hyperparameters.[20] In particular, for each algorithm, we test a maximum of 10 hyperparameter combinations. Given the randomized grid search, we actually deploy 200 models (more details in Table S2).[21]

Algorithmic Training. We start by dividing the original dataset into training and test sets, adopting a 70-30 rule, meaning that 70% of the observations were associated with the training set, while 30% were associated with the test set. The original training and test sets thus include 106,483 and 45,636 observations, respectively. Figure S2 plots the distribution of train and test sets per year. To address the imbalance issue, we have adopted two combined strategies. In the creation of the train and test set, we stratified the target variable such that the relative amount (or ratio) of positive and negative instances was constant in the two sets.[22] Then, we deployed

[18]The Roc-Auc score is calculated as follows: Roc-Auc $= \int_0^1 \text{True Positive Rate}\, d(\text{False Positive Rate})$. The integral in the formula represents the area under the ROC curve and the larger this area, the better the model's overall performance. The Roc-Auc score is particularly useful in situations where the class distribution is imbalanced, and it provides a robust evaluation metric that is insensitive to the specific threshold chosen for classification. It assesses the model's ability to rank positive instances higher than negative instances across different decision thresholds.

[19]Supplementary Materials section A.3 provides further details on each approach.

[20]Unlike internal parameters adjusted within the model during training, hyperparameters are predetermined before the training process and are commonly refined through methods such as trial-and-error, experimentation, or the application of optimization techniques. Examples of hyperparameters include the learning rate or the number of trees in a random forest.

[21]The total number of potential model candidates was 394 which, multiplied by four different sampling approaches (i.e., original, SMOTE, ADASYN, SMOTE+Tomek) equals a total of 1,576 model candidates. The entire predictive pipeline takes about 22 hours on an 8-core laptop and around 6 hours on a 32-core server.

[22]The stratification step broke the temporal order of the dataset, as we forced the train and test set to have the same ratio between infiltrated and non-infiltrated observations. We specifically assess the impact of this decision in Section 4.2.

the synthetic generation of observations solely in the training set via SMOTE, ADASYN, and SMOTE+Tomek, as anticipated above.

After the synthetic upsampling step, we searched the grid of selected hyperparameters for each algorithm and sampling approach to find the best architecture: the optimization criterion was based on maximizing out-of-sample recall (i.e., recall measured on the test set). Optimization has been performed via 5-fold stratified cross-validation which works by splitting the training set into 5 training subsets. The overall training set performance is computed as the average of the performance of each subset. Once cross-validation is concluded, the actual performance is evaluated on the test set (which, again, in our case amounts to the 30% observations in the original dataset).

4 Predicting infiltrated city councils

4.1 Predictive Performance and Model Explanation

As anticipated, we sample from a pool of 1,576 model candidates, comparing a total of 200 combinations of algorithmic approaches and architectures. Figure 2 visualizes the performance of the best-searched architecture for each algorithm and sampling approach in terms of recall (using the typical 0.50 probabilistic threshold). XGBoost trained on a dataset with ADASYN-based synthetic oversampling outperforms all other models, achieving a remarkable recall of 0.96 in the test set. In other words, XGBoost with ADASYN correctly predicts 96% of the actually infiltrated municipalities in the test set.[23]

The second-best model, also based on XGBoost, relies on SMOTE+Tomek as the synthetic upsampling technique, reaching a recall of 0.90.[24] Interestingly, XGBoost is the only algorithm reaching a considerably high recall without the use of any synthetic sampling approach, correctly identifying 72% of actually infiltrated municipalities.[25]

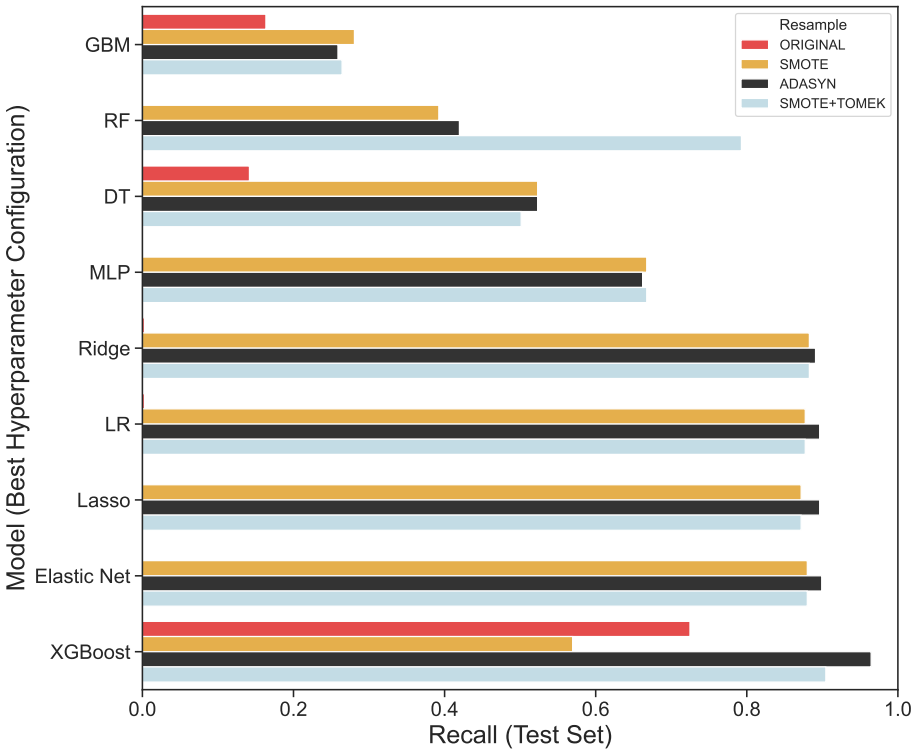
[23]The best model, i.e., XGBoost with ADASYN, has the following hyperparameter combination: 100 estimators, trees' max depth of 5, a learning rate of 0.1, $\gamma = 0$, and $\alpha = 0.5$, with a rescaling of weight given to the positive class equal to 50.

[24]This specific algorithm is composed of 50 estimators with a maximum depth of 10, learning rate equal to 0.1, re-weighting to the positive class of 50, and 0 values for both γ and α.

[25]For this model, Tables S3 reports the confusion matrices for the entire sample.

Figure S5 in the Supplementary Materials further elaborates on these performance results, complementing them with Precision and Roc Auc Scores on the test sets. First, the two best-performing models in terms of recall are also the best ones when considering the Roc-Auc score, achieving 0.92 and 0.90, respectively. Second, these two models perform low in terms of precision (0.06 and 0.07, respectively), indicating that the probabilistic 0.5 threshold may be too costly on the side of false positives. We tackle this potential issue in Section 4.3.[26]

FIGURE 2. Predictive Performance: Recall



Note: Recall is calculated using the typical 0.50 threshold and is computed on the test set, per each model and each sampling approach. Each bar refers to the best performing architecture emerged from the randomized grid search in terms of hyperparameter combination. Models are ordered from top to bottom based on increasing recall scores from the ADASYN resample. XGBoost with ADASYN sampling reaches the highest recall performance (0.96).

[26]To contextualize these findings, our best model performs comparably to or better than other predictive approaches in related studies. For example, Ash et al. (2020) achieved a maximum recall of 0.56 and a maximum Roc-Auc score of 0.78 in their study on corruption in Brazil. Colonnelli et al. (2020) investigated municipality-level corruption in Brazil and obtained a recall of 0.93 with their best model. In the analysis of politically connected firms in the Czech Republic, Titl et al. (2024) achieved a recall slightly exceeding 0.90 for their best model. Finally, de Blasio et al. (2022), focusing on corruption in Italian municipalities, reached a maximum recall of 0.74.

To gain insights into the key variables driving the prediction, we employ SHAP (SHapley Additive exPlanations, Lundberg and Lee (2017)).[27] SHAP is a computational approach designed to gain insights on how a machine learning model makes predictions. It does this by comparing the model's output when a specific feature is included versus when it is excluded. We use SHAP to explore how different groups of variables in our feature space contribute to the overall explanation (for more details, see Section A.4 of the Supplementary Materials).

Figure S6 illustrates the top-5 influential features for each group of variables. Public spending variables emerge as highly impactful: three of the top-5 predictive items relate to education. Specifically, higher current expenses allocated to education, both at time t and t-1, are associated with a lower risk of infiltration. Additionally, ideological and electoral variables also play significant roles. For example, municipalities ruled by left-wing or female mayors are less likely to be infiltrated, consistent with prior research linking corruption to gender (Brollo and Troiano, 2016) and political ideology (Hessami, 2012).

4.2 Sensitivity and Additional Tests

To evaluate the robustness of our predictive results, we perform three sets of distinct sensitivity tests and an additional analysis focusing on wrongful dismissals.

Random Seeds. First, we replicate our analyses testing 100 different combinations of random seeds. Specifically, we vary the random seeds used for the train and test split phase and the ADASYN upsampling phase.[28] Figure S7 displays the distribution of recall for the 100 models. The figure shows that recall is always higher than 0.90: on average, recall is 0.95, with identical median. The graph outlines how the performance of our main model is not simply the byproduct of an anomalous configuration of the train and test sets, and that our algorithm performs practically identically regardless of how the two are arranged. Figure S8 instead displays the

[27] SHAP, rooted in coalitional game theory (Shapley, 1952), reveals the most influential features in a model's prediction output both for local and global measures. While other methods like permutation-based feature importance exist to identify crucial features, SHAP focuses on features with a higher weight on predicted outcomes.

[28] The effect of random seeds is receiving increasing attention (Madhyastha and Jain, 2019; Beam et al., 2020), as varying random seeds can lead to dramatic variability in the predictive performance thus undermining the reliability, replicability and generalizability of results across various domains.

Pearson' r coefficients of the predicted risk of infiltration for each municipality-year observation in our dataset, across the 100 models. The plot shows that correlation across predicted risks is extremely high, remaining consistently over 90% over the entire matrix: Different models not only result in almost identical predictive performances, but also in very close measures of predicted risk.

Feature Selection. Second, we conduct sensitivity analyses acting on three dimensions: the removal of specific features, the use of different ratios between train and test sets and the preservation of the temporal structure of the data when splitting it into training and test sets, effectively treating the data as a panel dataset (e.g., using all municipalities from 2001 to 2014 in the train set, and testing on all municipalities from 2015 to 2020). This allows us to evaluate whether our index's performance depends on some specific features and if it is a byproduct of the stratified train-test split. Table S4 and S5 report the findings. The former table follows the stratification of the target variable (as in the main model), while the latter maintains the temporal structure when separating between train and test sets. The results indicate that retaining all features leads to improved overall performance. In all cases, the machine learning model performs better with a smaller test set size. Ideology does not significantly influence infiltration prediction. Removing public spending variables generally increases recall but it substantially decreases precision (approximately four times lower). Moreover, maintaining the temporal order of observations does not significantly reduce predictive performance.[29]

Sample Selection. Finally, Table S6 reports the results of five additional predictive exercises related to the selection – and construction – of our sample. In the first one, we train our model solely using municipalities from Southern regions, thus focusing on municipalities historically more prone to mafia presence. Overall, the prediction does not improve, as we observe slightly higher precision but lower recall. Second, we train the model using municipality-level dummy variables in the feature space (instead of regional features), aiming to assess the importance

[29] As we use lagged measurements for our public spending variable, there was a possibility of leakage of information when the same municipality at time t is in the train set and at time t-1 is in the test set (or vice versa). These sensitivity tests help excluding concerns on data leakage, as we keep the temporal structure of the dataset.

of the temporal variation in our predictive exercise. The performance remains very similar to the one commented in the main models, underlying the importance of temporal variation for predicting infiltration effectively. Third, we train the model only on municipalities dismissed for mafia, testing the algorithm's capability to differentiate temporal variation within infiltrated municipalities. We find comparable recall and higher precision, further illustrating the role of temporal variation. Fourth, we extend the previous configuration by excluding all years post-dismissal, following the methodology proposed by Ash et al. (2020). This approach aims to reduce the possible effect of post-dismissal institutional monitoring. The results are similar to the previous exercise. Fifth, inspired by Colonnelli et al. (2020), we construct a cross-sectional sample by aggregating all features into two-decade averages. This strategy assesses whether emphasizing time-fixed factors, such as regional dummies, while reducing temporal variation from public spending and electoral outcomes, affects performance. We find a substantial drop in recall (0.35 for the original sample, 0.69 for ADASYN), once again underscoring the importance of time-varying factors.

Wrongful Dismissals. In the Appendix (see Section A.5), we also exploit available information on dismissal decrees that are revoked (or nullified).[30] Decrees are generally revoked as a result of an appeal to the Regional Administrative Court that overturns the decision after the city council has already been dismissed. Therefore, as a complementary exercise, we trained machine learning models — mirroring the pipeline seen in the main analyses — to identify wrongful dismissals versus rightful dismissals. Therefore, in this exercise, we compare similar local governments that were all subjected to dismissal, with some later being reinstated.[31] Results in terms of recall are displayed in Figure S9. Our algorithm allows us to correctly identify 74% of all decrees that were nullified or revoked in the test set, indicating a high ability in discriminating between true versus wrongful dismissals, in the light of the very small sample size. This result is also relevant because in the first year after dismissal, municipalities report

[30] In our main analyses, when we predict mafia infiltration, we treat wrongful dismissals as 0, hence as non-infiltrated observations.

[31] The only difference compared to the main model is that, given the much lower number of observations, we employed feature selection to remove variables that were not particularly relevant for the prediction of wrongful dismissals, thus contrasting the potential issue of dimensionality. For this reason, we used Linear Support Vector Classification with $L1$ regularization equal to 5, reducing the number of features from 223 to 24.

a 45% drop in public investments (Galletta, 2017). Therefore, being able to predict whether one municipality is wrongfully dismissed can have positive economic effects both for the local government as well as other stakeholders, such as firms working with or for the public sector.[32]

4.3 A Synthetic Measure of Infiltration Risk

Description and Yearly Performance. As explained in Section 3.2, we chose to optimize recall as the main performance measure, while trying to avoid an excessive rate of false positives. To address this aspect, we examine our synthetic measure of mafia infiltration risk across various probabilistic thresholds in terms of both recall and precision.[33]

Figure 3 report these results within specific areas of the distribution (from above median to the top 100 observations). We display both the contemporaneous predicted risk and the lagged ones at t-1 and t-2.[34] The rationale is to assess whether – and to what extent – our measure has practical value in anticipating infiltration detection. Our focus is on yearly prediction, that is we compute the performance of the prediction for each separate year.

Starting from the bottom of each subplot, when we consider observations with a score above the median, we always obtain 100% recall and a very low precision. The very low precision, in this case, is a mechanical artifact of our sample which includes very few positive instances (0.83%).

We consider the 95$^{\text{th}}$ percentile threshold to be a very interesting one in the trade-off recall-precision. In this group, which includes approximately 360 municipalities every year, the percentage of truly infiltrated local governments is on average 14.6% (SD=4.01) (the percentage remains practically identical when using the t-1 and t-2 versions). Noteworthy, targeting this subset the recall remains very high, ranging from 81 to 97%, with an average of 88.9%. The

[32]Figure S10 provides explainability results, indicating which features mostly contribute to the prediction of wrongful dismissal.

[33]On average, the predicted risk of infiltration over the entire dataset is 0.15, while the median value is 0.02, indicating a right-skewed distribution. Figure S11 shows the high correlation between the trends of the synthetic measure in the train and test sets. Figure S12 in the Supplementary Information visualizes the distribution of the synthetic measure for the entire dataset, including observations from both the train and test sets. Figure S13 reports the distributions per each macro-region. Figure S14 displays the distribution of the predicted risk for the entire country in three years, i.e., 2001, 2010, 2020.

[34]We exclude precision and recall using the t-1 lagged version of the index for year 2001, as 2001 is our first year and the lagged does not exist, as well as 2002 for the 2-year lagged version, for the same reason.

t-1 lagged version offers a highly comparable performance, ranging from 78% to 93%, with an average of 84%. When using the t-2 risk to predict infiltration, recall is, on average 79% (min=71, max=87, SD=5.23), meaning that we are able to predict four out of five cases of true infiltration two years in advance. In other words, devoting investigative attention to around 4.5% of municipalities (per year) would uncover almost all truly infiltrated local governments, without incurring in an excessive rate of false positive.

If one further restricts the focus to the top 100 municipalities at risk, yearly percentages of infiltrated local governments range from 16% to 37%, with an average of 28.15 (17%-35% for the lagged at t-1, average 26.6; 16%-34% for the lagged at t-2, average 24.8). Recall decreases, as targeting the 100 local governments most at risk each year allows to identify between 31 to 70% of all municipalities that are found infiltrated by the Italian state, with an average of 48.5% (44.45% for the lagged at t-1, and 40.75% at t-2). This means that it is sufficient to monitor less than 1.3% of all municipalities in Italy to uncover around 50% of all truly infiltrated local governments, with a probability of finding an actually mafia-infiltrated municipality every four observations.

Overall, these results not only speak to the predictive performance of our approach for what concern contemporaneous infiltration: They demonstrate that the that predictive power remains very high even when predicting infiltration one or even two years in advance.[35]

Additional Tests. Figure S17 shows the average cumulative and marginal gains in recall in percentage terms, at the yearly level. Specifically, we investigate how many municipalities authorities should investigate yearly to reach a perfect recall. For this analysis, municipalities are grouped in sets of ten, ranked in descending order of predicted infiltration probability. We use the contemporaneous predicted risk, as well as the lagged predicted risks at t-1 and the lagged t-2 to measure performance. In terms of cumulative gain, the plot shows that, by only focusing on a very small fraction of Italian local governments, would lead to the detection of almost all infiltrated local governments. Concerning marginal gain, we indicate how the curve decreases reaching a null marginal gain around the threshold of 400 scrutinized local governments per

[35]In a related exercise, reported in Table S7 and discussed in Appendix Section S7, we report the performance of the index in the years 2021–2023.

year. This approach offers flexibility between recall and precision. Investigators can either monitor a small number of local governments for a high success rate or aim for comprehensive coverage of most infiltrated local governments, accepting the trade-off of more investigations and a higher incidence of false positives.

We additionally replicate these analyses using the average prediction as a result of 100 models run with different combinations of random seeds (see Paragraph 4.2) to further verify that our findings are not driven by mere chance. Figure S15 indicates that our results are robust to this sensitivity test, as all performances remain almost identical. Furthermore, we expand the analysis of the lags, up to 6 years (Figure S16). The performance deteriorates as the lag increases: To exemplify, recall goes from 89.2% for the observations with a predicted risk higher than the 95^{th} percentile to 61% at $t-6$.

Finally, to better understand the temporal variation in the data, we look at the prediction within the sample of positive instances, i.e. municipalities dismissed for mafia infiltration. Figure 4 clearly illustrates lower risk scores in the years preceding mafia infiltration. The trend shows a marked increase leading up to the infiltration and then after the dismissal, the risk decreases. This suggests that the dismissal of city councils might have a lasting positive impact on reducing the risk of mafia infiltration.[36] Figure S19 distinguishes between the year of the dismissal and the years right before in the same electoral term (infiltration years). The average predicted risks for infiltration and dismissal years are nearly identical: This provide additional evidence of how this index can well in advance predict where and when infiltration is likely to occur.

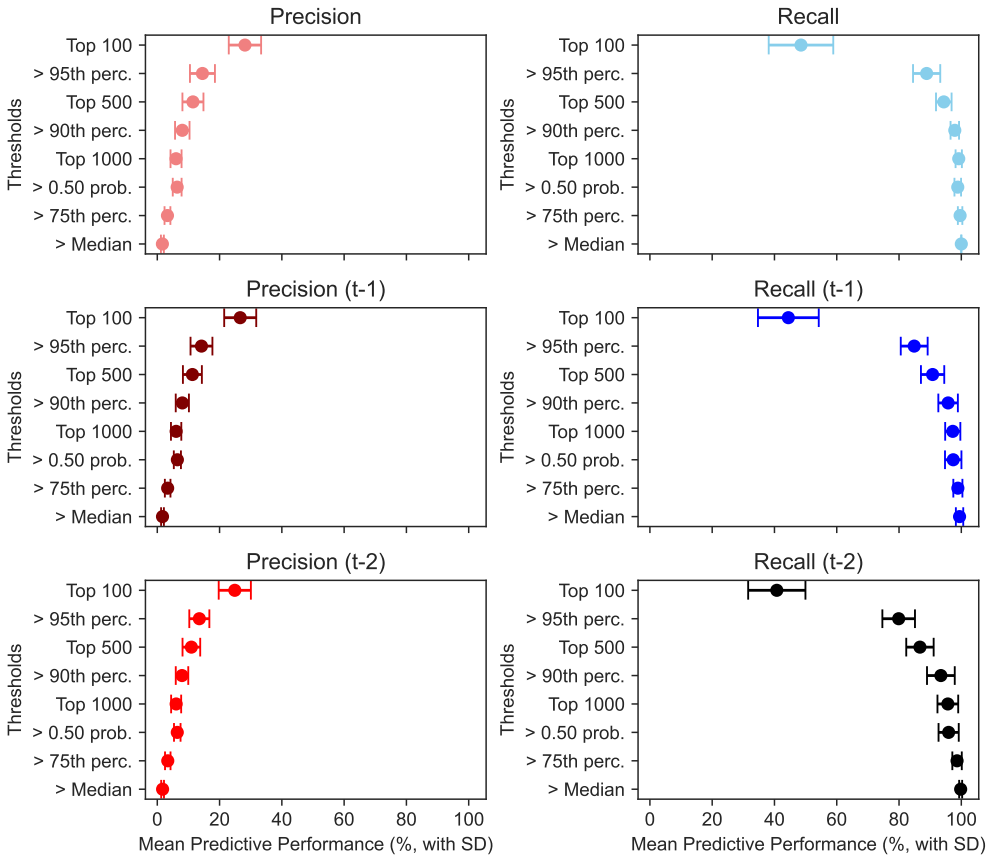
4.4 Validation and Cost-Benefit Considerations

As previously highlighted, false positives might include infiltrated local governments that have not been detected by the Italian state. While it is impossible to ascertain this definitively, we can provide indirect evidence supporting this interpretation.

First, consider detected infiltrated local governments. It is intuitive to expect red-flag events

[36]This descriptive analysis is supplemented by Figure S18, where we focus on municipalities dissolved in more than two non-consecutive electoral mandates. The predicted index remains very high both during the first and the second dismissal (red dots), while it often decreases between the two infiltration periods.

FIGURE 3. Precision and recall for various thresholds and rankings - by year (2001-2020)

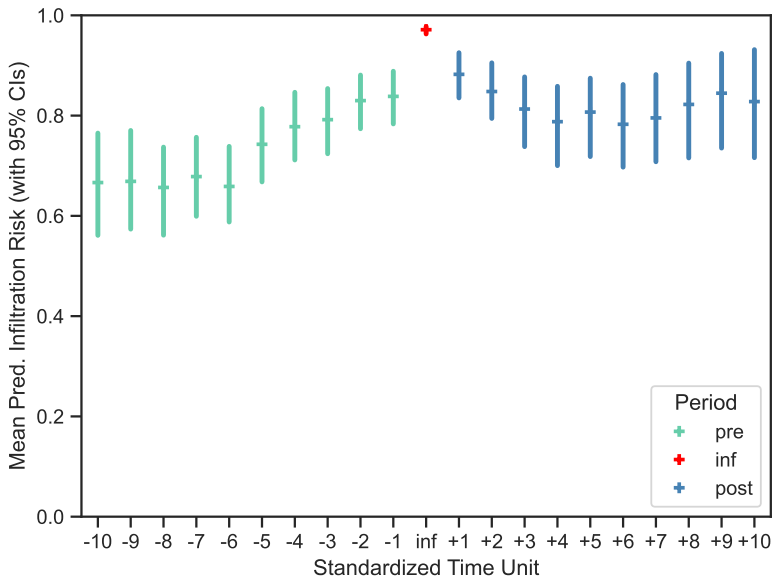


Note: We report mean and standard deviation. Precision and recall are shown as percentages. Precision (t-1) and Recall (t-1) refer to the performance when using the risk indicator at t - 1 to predict infiltration at t. Precision (t-2) and Recall (t-2) refer to the performance when using the risk indicator at t - 2 to predict infiltration at t. Performance is considered using predicted risk for observations both in the train and test sets. Threshold values are calculated by considering each yearly distribution of the predicted risk of infiltration.

related to organized crime in these areas, as such events could correlate with mafia influence on local politics. If false positives were also infiltrated, we would expect a similar association. Note that we are not claiming any causal effect but simply expect similar patterns between detected infiltrated governments and those suspected of infiltration according to our index.

To test this intuition, we collected data on three proxies of organized crime presence at the municipal level: (i) violent attacks against local politicians, available from 2010 to 2020 (Daniele and Dipoppa, 2017); (ii) mafia murders and attempted murders available from 2004

FIGURE 4. Mean Predicted Infiltration Risk Trends for Dissolved Municipalities



Note: The plot displays the distribution in the mean predicted risk of infiltration for municipalities that were dissolved at least once (n=175). For ease of interpretation, we collapse all time units where a municipality was infiltrated in a single point. Additionally, we censor the pre- and post-periods to -10 years and +10 years.

to 2013; (iii) assets seized from criminal organizations available from 2000 to 2013.[37]

We then compare local governments infiltrated by organized crime with those not infiltrated. Within the latter, we create a third group, i.e. a dummy variable equal to 1 for *Suspected* local governments, defined as those in the top 1% of our index. In other words, we create a dummy for the most at-risk false positives in our index.

Table 1 reports these panel estimates, controlling for municipality and year (or region-year) fixed effects. We observe two patterns: first, there is a strong association between infiltrated local governments and two of the three proxies of organized crime presence (attacks and murders); second, we find this association also for *Suspected* local governments. This association has a lower magnitude for attacks and murders, while it remains statistically significant also for seized assets. Overall, this suggests an interesting similarity between infiltrated local governments and the most at-risk false positive ones.

[37]Given the sparsity of these data, we coded them as dummy variables in our analysis, thereby focusing solely on the extensive margin. For example, municipalities-year observations with more than one mafia murder are only 0.1%.

TABLE 1. Local governments infiltrated and suspected - panel analysis with mafia red-flags

	(1) Attack	(2) Attack	(3) Mafia Murder	(4) Mafia Murder	(5) Seized Asset	(6) Seized Asset
Infiltrated	0.105***	0.0774***	0.0404***	0.0419***	-0.0168	-0.0156
	(0.0250)	(0.0254)	(0.0103)	(0.0104)	(0.0180)	(0.0182)
Suspected	0.0577***	0.0376***	0.0193***	0.0218***	0.0227**	0.0228**
	(0.0136)	(0.0138)	(0.00578)	(0.00579)	(0.00980)	(0.0100)
Mun. FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Region-Year FE	NO	YES	NO	YES	NO	YES
Observations	22,420	22,403	18,342	18,342	24,049	24,049
R-squared	0.506	0.523	0.545	0.564	0.480	0.495

Note: This table shows the results of a panel regression analysis where the units of observations are local governments (on average in power for five years). The analysis includes electoral year and municipality fixed as well as region-year fixed effects in some specifications. *Infiltrated* = 1 for local governments dismissed for organized crime; *Suspected* = 1 for local governments in which are index is above 0.99 at least for one year in a specific electoral term. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

A related consideration is the cost-effectiveness of investigating local governments at high risk among the false positives. The primary costs would involve the opportunity cost of police labor and ordinary equipment temporally devoted to these new investigations instead of other activities. While these costs are relatively limited – if this diversion is concentrated in space and time –, the benefits of uncovering and dismissing an additional infiltrated local government are substantial. According to our calculations based on Di Cataldo and Mastrorocco (2022), criminal organizations capture almost 1 million euros per year per municipality in expenditure, by manipulating construction and waste management expenditures.[38] Furthermore, based on Fenizia and Saggio (2023), we estimate additional gains of around 14 million euros per year per municipality in terms of increased economic activity and employment after the dismissal. Specifically, they report long-term benefits including a 15% increase in industrial real estate prices and a 16.9% increase in employment.[39]

[38] Di Cataldo and Mastrorocco (2022) find a 14% increase in construction and waste management spending due to infiltration. Our estimates are based on the average spending in these sectors among infiltrated municipalities.

[39] Our estimates for industrial real estate prices are based on the average area of industrial real estates sold annually in Southern Italian municipalities, using data from the Italian Tax Agency. Estimates for employment gains can be directly estimated from Fenizia and Saggio (2023).

5 EU Funds and Mafia Infiltration in Local Politics

After having validated this synthetic measure of mafia infiltration, in this final part, we show how it can be used to investigate the determinants of mafia infiltration. We specifically explore the effects of a positive shock in public spending, involving increased transfers from the European Union. As highlighted in Section 1, our index represents an ideal measure to investigate this question, as it is difficult to find alternative measures due to lack of local and time-varying data on mafia infiltration.[40]

5.1 The 2007-2013 Wave of EU Funds

Italy, being a significant beneficiary of European funds under the Regional Policy, the primary investment policy of the European Union, receives support through two major funds: the European Regional Development Fund (ERDF) and the Cohesion Fund (CF).[41] Implementation of this policy involves collaboration between national, regional, and municipal governments in partnership with the European Commission. Notably, six out of twenty Italian regions fall under the category of recipients, known as "Convergence regions", including Puglia, Basilicata, Calabria, Campania, Sardinia, and Sicily. These regions qualify for a much larger share of funding based on the 75% GDP rule, meaning regions with a GDP per capita below 75% of the EU average are eligible.

Here, we specifically focus on the EU funds wave in the years 2007–2013 for two reasons. First, in the period 2007–2013, compared to the previous wave (2000-2006), there was a considerable increase in the amount of transfers allocated to the "Convergence Regions". In the period 2007-2013, Italy received a total of 74 billion euros, with 56 billion allocated to the "Convergence Regions," including areas characterized by a substantial presence of organized crime.[42] This represents a considerable increase compared to the 30,7 billion euros allocated in

[40] Ideally we could use data about municipalities dismissed for mafia. However, no municipality in the treatment or control group has been dismissed for mafia after 2006 (specifically within the 100km border defined in the next paragraph).

[41] https://ec.europa.eu/regional_policy/en/policy/what/investment-policy/

[42] https://politichecoesione.governo.it/it/politica-di-coesione/la-programmazione-2007-2013/risorse-2007-2013/

the 2000-2006 wave.[43]

Second, the 2007-2013 wave underwent significant reforms by the EU Commission, aligning with the amendments introduced by the Lisbon Treaty. The most important aim of the reform was to further simplify and decentralize the process of programming and implementation. Notably, there was a shift towards increased delegation, with the administration of funds now being jointly managed by the Member States and the Commission. Furthermore, the management of projects saw a further decentralization to the local level (Davies and Polverari, 2011). This resulted in a greater portion of funds being directly managed and spent by local governments, with over 75% of funds, totaling 46 billion euros, falling under their purview (De Angelis et al., 2020). Thirdly, monitoring was scaled back, with mid-term project evaluations becoming optional (Manzella and Mendez, 2009).

Overall, this increase in public investments – if followed by economic growth – might reduce the fertile ground for the development of organized crime and its infiltration in politics, while lower monitoring and higher decentralization, might improve accountability, resulting in a higher effectiveness of public investments. Moreover, the enlargement of this program might signal the interest of national and supra-national institutions in investing in disadvantaged areas, decreasing the legitimacy of mafias in favor of formal institutions.

On the other hand, more funds could potentially catalyze mafia groups seeking to misappropriate these economic resources, thereby distorting their allocation at the municipal level. These dynamics could be enhanced by more decentralization, as this might foster rent-seeking, facilitating the infiltration of criminal organizations in the management of such funds.[44]

5.2 Identification

To discern the impact of EU funds (2007-2013) on mafia political infiltration, we employ a Difference-in-Discontinuities (Diff-in-Disc) methodology (Grembi et al., 2016; Eggers et al.,

[43] https://leg15.camera.it/cartellecomuni/leg14/RapportoAttivitaCommissioni/testi/05/05_cap25_sch01.htm

[44] In line with a higher risk of rent-seeking, we observe a significant rise in infractions in the management of EU funds during this period. The number of frauds reported to the EU by the Italian National Anti-EU Fraud Committee peaked at 828 in 2009 and 1027 in 2010 compared to 206 and 391 respectively in 2005 and in 2006 (Senato della Repubblica, 2018). Decentralization is also listed as one of the elements that increased the risk of corruption in public procurements within the European Union (van Ballegoiij and Zandstra, 2016).

2018; Butts, 2023). First, we leverage the geographic discontinuity in fund allocation between two eligible regions (Campania and Puglia) and two ineligible ones (Molise and Lazio). Second, we consider the temporal dimension, with the start of the 2007-2013 wave in 2007. Figure S20 shows the change in the amount of EU funds received by municipalities within treated and non-treated regions in the sample, comparing the periods 2000-2006 and 2007-2013.

The geographic Diff-in-Disc allows us to eliminate the influence of compound treatments occurring at the same administrative boundary, whether observable or not.

In details, we estimate the following Difference-in-Discontinuities specification:

$$\begin{aligned} \text{Infiltration}_{i,t} = \alpha + \beta \text{Treated}_i \times \text{Post2006}_t + \eta f(\text{Distance}_i) \times \text{Post2006}_t \\ + \gamma \text{Treated}_i \times f(\text{Distance}_i) \times \text{Post2006}_t + FE_i + FE_t + \varepsilon_{i,t}, \end{aligned} \tag{3}$$

where $Infiltration_{i,t}$ is the infiltration risk for municipality i in year t, $Post2006_t$ a dummy recording if year t is before or after 2006, $Treated_i$ indicating whether the municipality belongs to Puglia or Campania, $f(Distance)$, 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 a final one including all municipalities in a region. We use either linear or quadratic specifications.[45] We weight observations using a triangular kernel, FE_i and FE_t are municipality and year-fixed effects, while $\varepsilon_{i,t}$ are robust standard errors adjusted for serial correlation.[46]

The Diff-in-Disc method assumes that time-varying confounders do not change differently between treated and control groups over the pre- and post-treatment periods. As a result, any confounding effects are accounted for by exploiting the time variation in the treatment. Table S11 reports the balancing tests for potential treatment effects with regard to several time-varying characteristics at different distances from the border. Treated and control units are remarkably similar in how these characteristics vary due to the treatment when closer to

[45] An alternative specification for the Regression Discontinuity is the local randomization (Cattaneo et al., 2016). However, as noted by Butts (2023), such identification through randomization local to the cutoff is not applicable in the context of geographic Diff-in-Disc, as the method would presume that municipalities can be assigned randomly on either side of the regional border. By administrative construction, instead, municipalities are always on a given pre-determined side.

[46] When using a dependent variable that is generated (Battaglia et al., 2024), it is common practice to estimate standard errors using bootstrap methods; In the main tables the p-values are derived with standard errors bootstrapped 10,000 times.

the border, but still remain highly comparable when moving far away from it.

Finally, the assumption of non-sorting around the threshold does hold, as municipalities cannot alter their regional affiliation, at least in the short run. No municipality has undergone such a change throughout the entire observation period.

The remaining requisite assumption is the parallel trend assumption between treated and control municipalities. To this aim and to get more insights into the dynamics of the post-treatment effect we estimate an event study version of the previous specification, formalized as:

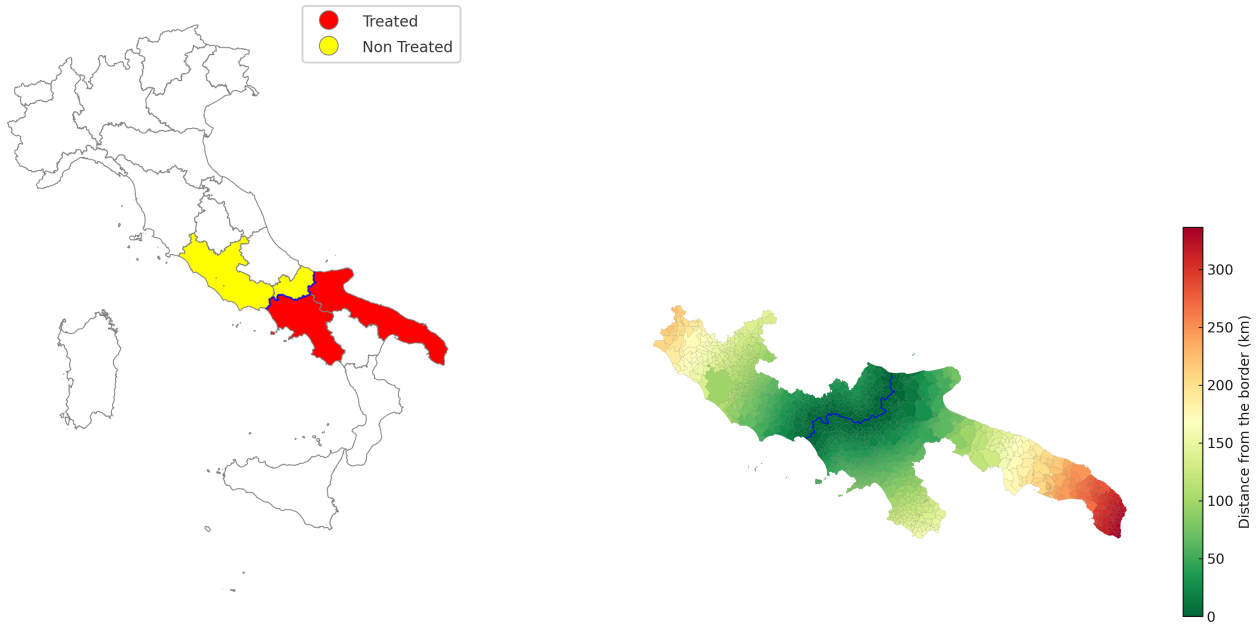
$$\begin{aligned}\text{Infiltration}_{i,t} = \alpha + \sum_{t=2001}^{2013} \beta \text{Treated}_i \times \text{Year}_t + \sum_{t=2001}^{2013} \eta_t f(\text{Distance}_i) \times \text{Year}_t \\ + \sum_{t=2001}^{2013} \gamma_t \text{Treated}_i \times f(\text{Distance}_i) \times \text{Year}_t + FE_i + FE_t + \varepsilon_{i,t}\end{aligned} \tag{4}$$

where we substitute the dummy $Post2006_t$ with a set of separate year dummies $Year_t$.

Figure 5 provides geospatial visualizations of the control and treatment regions in our analysis. In panel (a), we highlight in yellow Lazio and Molise, the two regions used as controls, while we color in red the treated ones: Campania and Puglia. In panel (b), instead, we specifically focus on these four regions by plotting the map at the municipal level, highlighting the border (in blue) separating treated and control and coloring each municipality based on its distance from the border. Figure 6 shows the average change in the predicted index before/after 2007. The map highlights how the wide majority of the municipalities experiencing sizable increases in the predicted index of infiltration lay below the red border in the treatment group.

Non-Classical Measurement Error. A potential limitation of our approach stems from the nature of our dependent variable, deriving from our predictive model of mafia infiltration. This model is trained on a non-random sample of detected infiltrated municipalities, introducing potential biases into our analysis (for instance, in terms of detection efforts changing over time or space). Specifically, this non-randomness may lead to a non-classical measurement error when using this measure as a dependent variable.

FIGURE 5. Geographical Distribution of Treated and Control Regions and Municipalities

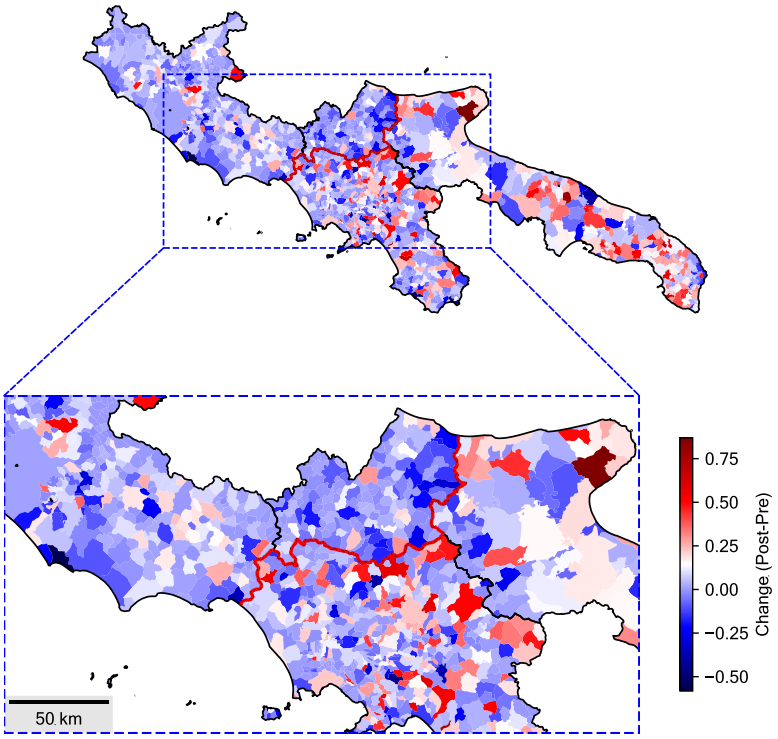


Note: On the left, we display treated and control regions; on the right, we report the treated and control regions, rendered with municipal boundaries and the border (in blue) separating the treated and control regions. The colormap quantifies the distance of a given municipality from the border, with green entailing lower distance and red higher distance.

In this setting, the relevance of this concern is very much narrowed. First, the Diff-in-Disc approach allows us to rule out the impact of time-invariant components of such a non-classical measurement error, including potential predispositions of law enforcement agencies to investigate specific areas. Second, the focus on municipalities near the border implies a high degree of comparability between treated and control municipalities in terms of time-variant characteristics, either observable or not, and so they are likely to be affected similarly by different shocks over time (see Table S11).

Third, by employing an event study approach, we address potentially time-varying biases, if they are not perfectly correlated with the timing and spatial dimensions of the shock. A possible violation of this statement could concern changes in detection contemporaneous to the 2007-2013 EU funds wave. For instance, one might in principle hypothesize that, following the reception of EU funds, treated regions are subjected to higher scrutiny by the state, thus implying increasing efforts in uncovering ties between mafia and politics. In a robustness test,

FIGURE 6. Difference of Predicted Risk of Infiltration between Pre- and Post-Treatment Periods



Note: The top subplot shows the difference between the average predicted risk in the 2001-2006 period (pre-treatment) and the average predicted risk in the 2007-2013 period (post-treatment) in the four regions under analyses. The bottom subplot provides a detail of the municipalities close to the shared borders between treated and control regions. Missing municipalities indicate missing predicted infiltration risk due to lack of electoral data.

we specifically address this concern, by testing the results on a predictive model that exclude all infiltrated observations in the 2007-2013 period. The rationale is that by removing possible byproducts of increased efforts (i.e., detected infiltrated municipalities in the time window under scrutiny), we should capture the sole effect of EU funds on the risk of mafia infiltration.

A second concern pertains to the direct impact of EU funds on the municipal budget, increasing available funds for capital expenditures. In principle, it could be argued that the influence of EU funds on the budget mirrors the impact of organized crime. Yet, there are several aspects that logically and institutionally contrast such hypothesis. Firstly, our features do not use crude amounts of Euros spent per each item, but rather consider the item-wise

composition of current expenses, capital expenses and the relation between current and capital expenses. Second, EU transfers first enter the budget in the revenue side, as they are transfers from a higher level of government, while our prediction framework only relies on spending categories. Nonetheless, to dispel any doubt regarding this hypothetical issue, we perform four robustness tests described in detail below.

5.3 Baseline Results

Table 2 reports the estimates for Equation 3. We evaluate twelve specifications, entailing two different polynomials (i.e., linear and quadratic) and six different geographic thresholds as distances to the borders. The coefficients clearly signal that the reception of EU funds in treated regions has a sizable and positive effect on the predicted risk of mafia infiltration: The coefficients show a magnitude increase of around 10-14 percentage points, with the only exception being the regional comparison (5-8 p.p. in columns 11-12). The latter comparison is the least reliable as it includes also municipalities considerably far from the regional border.

In Figure 7 we show the estimates of Equation 4, using both linear and quadratic specifications for all the different geographic thresholds. In the pre-treatment period, treated and control municipalities are on parallel trends. Point estimates are mostly close to zero and practically never statistically distinguishable from that. However, once EU funds are administered to the treated municipalities (i.e., those municipalities in Puglia and Campania) we uncover significant upward trends in the point estimate, signaling an increased predicted risk of infiltration. This finding, although characterized by different magnitudes, emerges at each distance threshold. Except for the broadest distance (i.e., the regional model), the linear and quadratic specifications lead to substantially identical estimates.

5.4 Robustness Tests

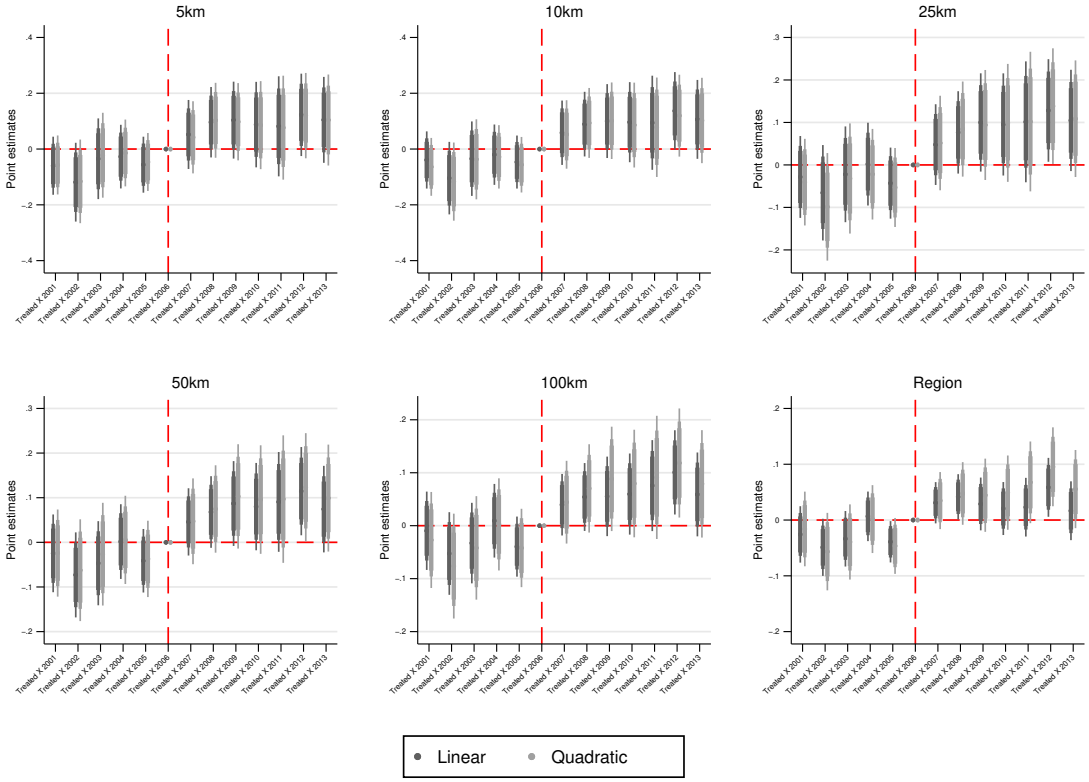
Robust coefficients. A wide set of robustness and additional models support our main results. First, replicating the main models with robust coefficients as proposed by Calonico et al. (2014, 2018, 2019, 2020), results remain practically identical (Table S8 reports the coefficients in tabular format, while Figure S21 shows the event study).

TABLE 2. 2007-2003 EU Funds and Risk of Mafia Infiltration

	(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	0.142***	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

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3. The dependent variable is the predicted risk of infiltration. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the post-treatment period. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

FIGURE 7. EU Funds and Risk of Mafia Infiltration – Event study



Note: This plot shows the effect over time of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 4. The dependent variable is the predicted risk of infiltration at the municipal level. Treated × 2006 indicates the year prior to the arrival of the EU funding and the red vertical dashed line separates graphically the pre- and post-treatment periods. Each subplot focuses on different geographical distances from the border, from 5km to the entire region. The plot reports each point estimate for the effect of EU funds on the risk of infiltration in a specific year along with two different polynomial specifications (i.e., linear and quadratic) along with 90%, 95%, and 99% confidence intervals.

Random Seeds. Second, we employ as measure of infiltration risk the average predicted risk obtained as a result of the 100 models deployed for sensitivity purposes using different combinations of random seeds (see Section 4.2). The outcomes of this robustness test are reported in Table S9 and Figure S22. The main effects are confirmed, holding for all the different specifications.

Binarized Infiltration Risk. Third, we use binary measures of infiltration risk discretizing the continuous indicator in deciles before the public spending shock in 2007. The effect has to be interpreted as the increase (or decrease) in the number of municipalities belonging to a given specific threshold-based category after the EU funds are received. The effects are larger when discretizing the variable using the 50% probability as the threshold (Figure S23). Regardless of the distance to the border and specification, we estimate an increase in the probability of mafia infiltration risk. The size of the estimates decreases as the binary indicator becomes stricter, but remains largely significant even when considering only municipalities with an 80% or higher probability of infiltration before 2007. In other words, the models indicate that the funds mostly affect those municipalities having a medium-high risk while maintaining virtually unchanged the number of those municipalities having extremely high risks (90% or higher). This result suggests that the funds have an attractive power to municipalities that were not previously subjected to intense mafia presence.

Placebo: Fake Borders. Fourth, we perform a placebo test by setting fake borders (Figure S24). More precisely, we define control and treated municipalities using artificial borders positioned at varying distances from the real border. Then, we replicate the estimation strategy using these simulated borders to define control and treated units. The effects are significant only when considering the true borders. This pattern is confirmed for all the threshold distances, from 5 km to the regional.

Removing Infiltration (2007-2013). Fifth, to contrast the objection that increased risk after EU funds reception might be driven by increased detection efforts in treated regions, we train a model removing all infiltrated municipality-year observations in the 2007-2013 period,

and then use such predicted risk to assess whether our findings still hold. Results are provided in Table S10, showing how the estimated coefficients remain sizable and highly significant, except for the regional models which, however, compare municipalities that present qualitative differences.

5.5 An alternative measure of mafia presence

A possible critique to this approach is the complete reliance on our predicted measure of mafia infiltration. Here, we validate our findings by using an alternative source of data, i.e. mafia-crime related investigations, which are available up to 2013 and represent a generic measure of mafia presence, rather than a specific measure of mafia infiltration in politics. Information on mafia crimes is based on the Italian Investigation System, a data collection system managed by the Ministry of the Interior. The system records details on investigation procedures authorized by the judiciary and carried out by police forces and allows us to compute the number of initiated procedures by municipality, year, and type of alleged offense.[47] A second limitation is the sparsity of this measure, as there are on average very few charges per municipality. To increase the power of this analysis we also include in our definition of mafia-related crimes offense categories that are not explicitly linked to mafias' actions but are commonly considered strong markers of their activity.[48]

We then replicate the Difference-in-Discontinuities approach. Results are significant for most distance thresholds (Table S16). By focusing on the event study, we noted a clear upward trend in the yearly estimates peaking in 2009 (Figure S25). Hence, we divided the post-treatment window into two different periods. This way, we identify larger, more consistently significant effects of EU funds on mafia-related crimes from 2009 on, while in the first two years after the reception of funds, no significant effect is discernible (Table S17). This finding is particularly relevant for two main reasons. On the one hand, it confirms that EU funds not only affect our indicator but have direct implications also on existing measures of mafia. On the other hand,

[47] Unfortunately, we have access to detailed criminal records at the municipal level only for a shorter time window than in the baseline analysis, that is the period 2004-2013.

[48] More specifically our measure of mafia-related crimes includes: threats, extortion, criminal or mafia association, mafia murder, money laundering and arsons.

the fact that mafia-related crimes increase with a delay of one/two years, while our indicator is immediately impacted, suggests that we might anticipate mafia presence compared to standard detection measures.[49]

5.6 Direct impact of EU funds on Municipal budgets

Pre- and Post- Distributions of Public Spending Features. As outlined earlier, a potential criticism of our identification strategy is that our predictive model relies on a wide array of public spending variables. In turn, fluctuations in spending patterns influenced by the receipt of EU funds could lead to higher predicted risks of infiltration.

First, we examine a Placebo region, i.e., Sardinia. While Sardinia was eligible for the 2007-2013 EU funds wave, due to historical and geographical criteria, it is an area with minimal presence of criminal organizations. Hence, in Sardinia, we do not expect any change in the predicted presence of mafia in politics following the arrival of EU funds.

We start by looking how the distribution of public spending variables changes in the pre and post-treatment. Specifically, if and how they change in the 2000-2006 window compared to the 2007-2013. We have computed Mann-Whitney U tests for significance across treated regions (Puglia and Campania) and Sardinia, that received funds in the 2007-2013 but did not record any city council dismissal over the 1991-2020 period.[50] The underlying intuition is that public spending variables should change their distribution only in treated regions (as a result of mafia infiltration) and not in the placebo. The results (Table S12) support this hypothesis: 40% of all public spending variables show a significant change in Puglia and Campania after the treatment (i.e., the reception of the 2007-2013 wave of EU funds). Conversely, public spending in Sardinia's municipalities has not changed in terms of overall distribution from the pre- to the post-treatment period. This analysis indicates that the allocation of municipal resources by local councils does not exhibit similar patterns between mafia presence and the use of EU funds.

[49]We also thoroughly verified that there is no other policy change at the regional border discontinuously changing in 2008 or 2009, which might represent an alternative explanation for the increase in mafia charges in treated areas.

[50]For completeness, two municipal councils were initially dissolved in Sardinia: Baratili San Pietro and Santu Lussurgiu in 2019. Both decrees, however, were then revoked.

Placebo Test: Sardinia. Second, we employ a straightforward Difference-in-Differences model, comparing municipalities in Sardinia with those in the control regions before and after the start of the funding wave 2007-2013.[51] As expected, we do not observe an increase in mafia infiltration; rather, there is a decline in the predicted index (Table S13).

Removing Capital Spending Features. As third strategy, we exploit the fact that EU funds can be spent by local councils only for investments (e.g. in local infrastructures, digital technologies, local transports), which are computed in the capital spending part of the budget. To exclude the direct impact of EU funds on our index, we re-compute the prediction exercise after excluding all capital spending variables. The recalibrated index maintains similar predictive accuracy, with the recalculated recall rate being consistent (0.97). This adjusted index then serves as the dependent variable in our Difference-in-Discontinuities analysis. The results are reported in Figure S26 and Table S14. The results are very similar to our main findings, discarding the idea that the arrival of EU funds is directly having an effect on the index of mafia infiltration in politics.

Removing Contemporaneous Public Spending Features. Fourth, we performed a test training our model removing contemporaneous public spending features. The rationale was to verify that our risk of infiltration was not mechanically determined by contemporaneous injections of EU funds. Previous findings are confirmed: Figure S27 shows the event study results using the indicator trained without contemporaneous public spending features, Table S15 displays the diff-in-disc results.

Less Decentralization and Monitoring As a final test, we explore the heterogeneous effects between the two treated regions, which were differently exposed to the EU funds shock. The 2007-2013 wave allocated significantly more resources to treated regions, potentially amplifying their attractiveness to criminal organizations. However, the 2007-2013 wave of EU funds also delegated more powers in terms of monitoring and management to national, regional, and

[51] It is important to note that in this case we cannot integrate the difference-in-differences identification with a regression discontinuity to obtain a difference-in-discontinuities design since Sardinia and the control regions lack a shared geographical border

local governments: this increased decentralization might foster rent-seeking independent of the funding increase. By focusing on this mechanism, we can discard the idea that EU funds are mechanically determining the observed effects, as their impact is muted in this case.

To disentangle these two mechanisms, we exploit the fact that the control region Lazio also saw a mild increase in funds under the 2007-2013 wave compared to the previous wave. This increase is comparable in relative terms to that experienced by the treated region Campania (around 8% for Lazio and 4% for Campania), though definitively lower in absolute terms (14 Euros per capita increase in Lazio versus 54 Euros per capita in Campania).[52]

This implies that if we only compare municipalities in Lazio with those in Campania closer to Lazio than Molise, decentralization should be the prevailing effect. Conversely, when comparing municipalities treated municipalities closer to Molise than to Lazio, both mechanisms-decentralization and increased funding-are at work.

In Table S18, we focus on the first case, considering only municipalities in Lazio and Campania. We find substantial effects, although the estimates remain noisy due to the small sample size. In Table S19, we test the second scenario by comparing municipalities in Puglia and Molise. In this case, we also find significant effects, suggesting that both mechanisms are at work, and that decentralization and reduced monitoring alone might be sufficient to trigger increased organized crime infiltration in local politics.[53] Overall, this test provides further evidence against the idea that the increase in EU funds might mechanically determine an increased risk of infiltration.

6 Conclusions

In this study, we have introduced a synthetic measure of mafia infiltration at the municipal level in Italy. Employing a comprehensive machine learning framework, we design our mea-

[52]Specifically, Campania received 1,122 Euros per capita in the 2000-2006 wave and 1,176 Euros per capita in the 2007-2013 wave; Lazio received 171 Euros per capita in the 2000-2006 wave and 185 Euros per capita in the 2007-2013 wave. In the same period, Molise received fewer funds (-32%), while Puglia increased from 929 to 1,382 Euros per capita (+48%).

[53]Ideally, we could examine the intensity margin by considering the increase in EU funds within each municipality before and after 2007. However, this approach would be endogenous, as funding allocations result from negotiations between regional and municipal governments. Additionally, this analysis is not feasible because municipal-level data on EU funds are unavailable prior to 2007.

sure by merging data on public spending and electoral outcomes spanning from 2001 to 2020. Our methodology exhibits robust predictive power, achieving an out-of-sample recall of 0.96, accurately identifying (up to two years earlier) municipalities previously detected by Italian authorities as infiltrated. Furthermore, we emphasize the importance of our measure in directing investigative focus towards high-risk municipalities, serving as a potential tool to counteract mafia's influence into local politics. This research advances the literature on organized crime by providing a granular measure of mafia infiltration, which can be used by both scholars and policymakers.

Applying this measure to study the causes of this phenomenon, we focus on the impact of a public spending shock, i.e. an increase in transfers from the European Union. Employing a geographic difference-in-discontinuities design, we compare municipalities in treated (Campania and Puglia) and control regions (Molise and Lazio). Our results underscore a significant and substantial increase in infiltration risk associated with the reception of EU funds. Treated municipalities within 5 km of the border experience an increased risk of around 13-14 p.p. These results highlight how transfers might have unintended effects attracting organized crime towards new political arenas.

Beyond its academic significance, this work holds practical and policy implications. The increasing use of data-driven solutions for countering criminal phenomena has gained momentum in many countries. However, many such solutions often lack validation or sufficient predictive power. Our framework equips law enforcement and anti-mafia authorities with a set of tools to tailor investigative efforts, mitigating costs and enhancing their ability to counteract the mafia's persistent influence on politics. We believe this approach can be extended to other countries facing organized crime issues, such as those in Latin America, the Balkans, or Eastern Europe, whether in contexts of high violence or similar to Italy's, where corruption dynamics prevail.

References

Acemoglu, Daron, Giuseppe De Feo, and Giacomo Davide De Luca, "Weak States: Causes and Consequences of the Sicilian Mafia," *The Review of Economic Studies*, March 2020, *87* (2), 537–581.

_ , James A Robinson, and Rafael J Santos, "The monopoly of violence: Evidence from Colombia," *Journal of the European Economic Association*, 2013, *11* (suppl_1), 5–44.

Alesina, Alberto, Salvatore Piccolo, and Paolo Pinotti, "Organized Crime, Violence, and Politics," *The Review of Economic Studies*, March 2019, *86* (2), 457–499.

Angelis, Ilaria De, Guido de Blasio, and Lucia Rizzica, "Lost in corruption. Evidence from EU funding to Southern Italy," *Italian Economic Journal*, 2020, *6*, 355–377.

Ash, Elliott, Sergio Galletta, and Tommaso Giommoni, "A Machine Learning Approach to Analyzing Corruption in Local Public Finances," 2020.

Baraldi, Anna Laura, Giovanni Immordino, and Marco Stimolo, "Self-selecting candidates or compelling voters: How organized crime affects political selection," *European Journal of Political Economy*, 2022, *71*, 102133.

Barone, Guglielmo and Gaia Narciso, "Organized crime and business subsidies: Where does the money go?," *Journal of Urban Economics*, March 2015, *86*, 98–110.

Batista, Gustavo E. A. P. A., Ronaldo C. Prati, and Maria Carolina Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD Explorations Newsletter*, June 2004, *6* (1), 20–29.

Battaglia, Laura, Timothy Christensen, Stephen Hansen, and Szymon Sacher, "Inference for Regression with Variables Generated from Unstructured Data," *arXiv preprint arXiv:2402.15585*, 2024.

Battiston, Giacomo, Gianmarco Daniele, Marco Le Moglie, and Paolo Pinotti, "Fueling organized crime: The mexican war on drugs and oil thefts," 2022.

Beam, Andrew L., Arjun K. Manrai, and Marzyeh Ghassemi, "Challenges to the Reproducibility of Machine Learning Models in Health Care," *JAMA*, January 2020, *323* (4), 305.

Becker, Sascha O, Peter H Egger, and Maximilian Von Ehrlich, "Going NUTS: The effect of EU Structural Funds on regional performance," *Journal of Public Economics*, 2010, *94* (9-10), 578–590.

Blattman, Christopher, Gustavo Duncan, Benjamin Lessing, and Santiago Tobón, "Gang rule: Understanding and countering criminal governance," Technical Report, National Bureau of Economic Research 2021.

Brollo, Fernanda and Ugo Troiano, "What happens when a woman wins an election? Evidence from close races in Brazil," *Journal of Development Economics*, 2016, *122*, 28–45.

Buonanno, Paolo, Ruben Durante, Giovanni Prarolo, and Paolo Vanin, "Poor institutions, rich mines: Resource curse in the origins of the sicilian mafia," *The Economic Journal*, 2015, *125* (586), F175–F202.

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

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell, "On the effect of bias estimation on coverage accuracy in nonparametric inference," *Journal of the American Statistical Association*, 2018, *113* (522), 767–779.

_ , _ , and _ , "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs," *The Econometrics Journal*, 2020, *23* (2), 192–210.

_ , Matias D. Cattaneo, and Rocio Titiunik, "ROBUST NONPARAMETRIC CONFIDENCE INTERVALS FOR REGRESSION-DISCONTINUITY DESIGNS," *Econometrica*, 2014, *82* (6), 2295–2326.

_ , **Matias D Cattaneo, Max H Farrell, and Rocio Titiunik**, "Regression discontinuity designs using covariates," *Review of Economics and Statistics*, 2019, *101* (3), 442–451.

Campedelli, Gian Maria, "Explainable machine learning for predicting homicide clearance in the United States," *Journal of Criminal Justice*, March 2022, *79*, 101898.

_ , *Machine Learning for Criminology and Crime Research: At the Crossroads*, Routledge, June 2022.

Cataldo, Marco Di and Nicola Mastrorocco, "Organized Crime, Captured Politicians, and the Allocation of Public Resources," *The Journal of Law, Economics, and Organization*, November 2022, *38* (3), 774–839.

Cattaneo, Matias D., RocÃo Titiunik, and Gonzalo Vazquez-Bare, "Inference in Regression Discontinuity Designs under Local Randomization," *The Stata Journal: Promoting communications on statistics and Stata*, June 2016, *16* (2), 331–367.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, June 2002, *16*, 321–357.

Colonnelli, Emanuele, Jorge A. Gallego, and Mounu Prem, "What Predicts Corruption?," *SSRN Electronic Journal*, 2020. Available at SSRN 3330651.

Cressey, Donald R., "Methodological Problems in the Study of Organized Crime as a Social Problem," *The Annals of the American Academy of Political and Social Science*, 1967, *374*, 101–112. Publisher: [Sage Publications, Inc., American Academy of Political and Social Science].

Daniele, Gianmarco and Benny Geys, "Organised Crime, Institutions and Political Quality: Empirical Evidence from Italian Municipalities," *The Economic Journal*, 2015, *125* (586), F233–F255. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12237.

_ **and Gemma Dipoppa**, "Mafia, elections and violence against politicians," *Journal of Public Economics*, October 2017, *154*, 10–33.

_ **and** _ , "Fighting Organized Crime by Targeting their Revenue: Screening, Mafias, and Public Funds," *The Journal of Law, Economics, and Organization*, November 2023, *39* (3), 722–746.

Davies, Sara and Laura Polverari, "Financial Accountability and European Union Cohesion Policy," *Regional Studies*, May 2011, *45* (5), 695–706. Publisher: Routledge _eprint: https://doi.org/10.1080/00343404.2010.529118.

de Blasio, Guido, Alessio D'Ignazio, and Marco Letta, "Gotham city. Predicting corrupted municipalities with machine learning," *Technological Forecasting and Social Change*, November 2022, *184*, 122016.

Dell, Melissa, "Trafficking networks and the Mexican drug war," *American Economic Review*, 2015, *105* (6), 1738–1779.

della Repubblica Senato della Repubblica, Senato, "At Europe's Expense Community funds are defrauded in a thousand different ways. How does this happen? And whoâs combating this fraud?," Technical Report, Impact Assesment Office, Senato della Repubblica, Italy 2018.

Dube, Oeindrila and Juan F Vargas, "Commodity price shocks and civil conflict: Evidence from Colombia," *Review of Economic studies*, 2013, *80* (4), 1384–1421.

Eggers, Andrew C., Ronny Freier, Veronica Grembi, and Tommaso Nannicini, "Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions," *American Journal of Political Science*, 2018, *62* (1), 210–229.

Fenizia, Alessandra and Raffaele Saggio, "Organized Crime and Economic Growth: Evidence from Municipalities Infiltrated by the Mafia," December 2023.

Galletta, Sergio, "Law enforcement, municipal budgets and spillover effects: Evidence from a quasi-experiment in Italy," *Journal of Urban Economics*, September 2017, *101*, 90–105.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, July 2016, *8* (3), 1–30.

Haibo He, Yang Bai, Edwardo A. Garcia, and Shutao Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in "2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)" IEEE Hong Kong, China June 2008, pp. 1322–1328.

Heller, Sara B., Benjamin Jakubowski, Zubin Jelveh, and Max Kapustin, "Machine Learning Can Predict Shooting Victimization Well Enough to Help Prevent It," June 2022.

Hessami, Zohal, "On the link between government ideology and corruption in the public sector," in "ISNIE Conference Paper" 2012.

Hofman, Jake M., Duncan J. Watts, Susan Athey, Filiz Garip, Thomas L. Griffiths, Jon Kleinberg, Helen Margetts, Sendhil Mullainathan, Matthew J. Salganik, Simine Vazire, Alessandro Vespignani, and Tal Yarkoni, "Integrating explanation and prediction in computational social science," *Nature*, July 2021, *595* (7866), 181–188.

Index, GITOC Global Organised Crime, "Global Organised Crime Index - 2021," Technical Report 2021.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan, "Human Decisions and Machine Predictions*," *The Quarterly Journal of Economics*, February 2018, *133* (1), 237–293.

Limodio, Nicola, "Terrorism financing, recruitment, and attacks," *Econometrica*, 2022, *90* (4), 1711–1742.

Lundberg, Scott M. and Su-In Lee, "A Unified Approach to Interpreting Model Predictions," November 2017. arXiv:1705.07874 [cs, stat].

Madhyastha, Pranava and Rishabh Jain, "On Model Stability as a Function of Random Seed," September 2019. arXiv:1909.10447 [cs, stat].

Manzella, Gian Paolo and Carlos Mendez, "The turning points of EU cohesion policy, Working Paper Report to Barca Report," 2009.

Marcolongo, Giovanna, "Organized Crime, Earthquakes and Local Public Procurement," June 2023.

Mirenda, Litterio, Sauro Mocetti, and Lucia Rizzica, "The economic effects of mafia: Firm level evidence," *American Economic Review*, 2022, *112* (8), 2748–2773.

Moglie, Marco Le and Giuseppe Sorrenti, "Revealing âMafia Inc.â? Financial Crisis, Organized Crime, and the Birth of New Enterprises," *The Review of Economics and Statistics*, January 2022, *104* (1), 142–156.

Pinotti, Paolo, "The Economic Costs of Organised Crime: Evidence from Southern Italy," *The Economic Journal*, 2015, *125* (586), F203–F232. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12235.

Shapley, Lloyd S., "A Value for N-Person Games," Technical Report, RAND Corporation March 1952.

Sviatschi, Maria Micaela, "Making a narco: Childhood exposure to illegal labor markets and criminal life paths," *Econometrica*, 2022, *90* (4), 1835–1878.

_ , "Spreading gangs: Exporting US criminal capital to El Salvador," *American Economic Review*, 2022, *112* (6), 1985–2024.

Titl, Vitezslav, Deni Mazrekaj, and Fritz Schiltz, "Identifying Politically Connected Firms: A Machine Learning Approach*," *Oxford Bulletin of Economics and Statistics*, 2024, *86* (1), 137–155. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12586.

United Nations Office on Drugs and Crime, "Legislative Guide for the Implementation of the United Nations Conventions against Transnational Organized Crime," 2016.

van Ballegoiij, Wouter and Thomas Zandstra, "Organised Crime and Corruption - Cost of Non-Europe Report," Technical Report, European Parliamentary Research Service 2016.

SUPPLEMENTARY INFORMATION FOR:

MAFIA, POLITICS AND MACHINE PREDICTIONS

Campedelli G.M., Daniele G., Le Moglie M.

Table of Contents

A Predicting and Explaining Mafia Infiltration **47**

A.1 Description of the dataset 47

A.2 Reducing Class Imbalance 50

A.3 Tested Algorithms and Hyperparameter Grid Search 51

A.4 Explainability via SHAP: Technical Details 57

A.5 Supplementary Predictive Results 58

A.6 A Synthetic Measure of Infiltration Risk – Further Details 69

B EU Funds and Mafia Infiltration **81**

B.1 Additional Figures 81

B.2 Additional Tables 89

A Predicting and Explaining Mafia Infiltration

A.1 Description of the dataset

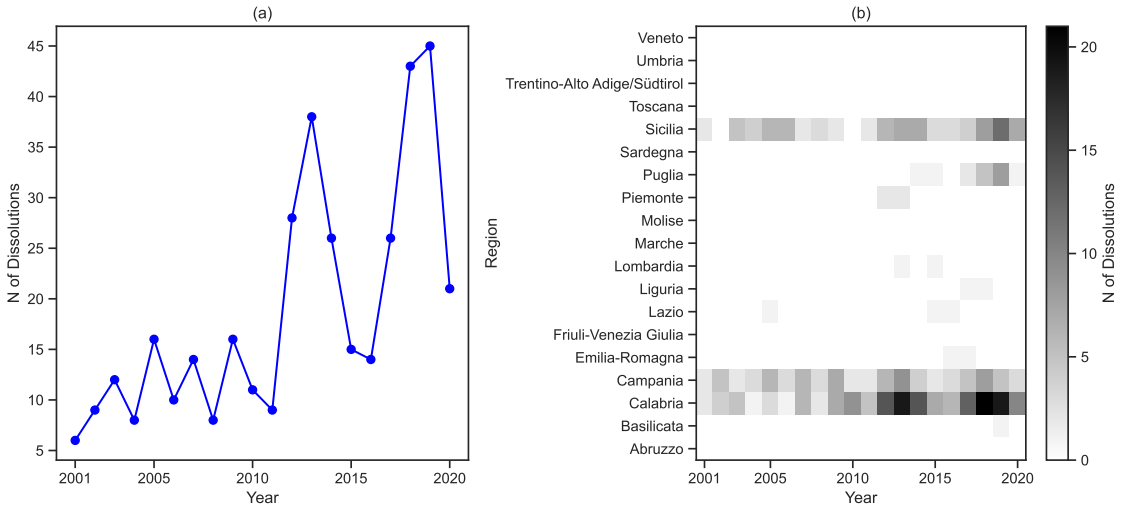
Figure S1 first displays the count of dismissal decrees by year at the national level (panel (a)) and the count broke down by region (panel (b)). Overall, the trend increases from 2010 to 2020, although non-monotonically, with substantial variation. The plot clearly shows the concentration of decrees in Sicily, Calabria, and Campania, along with Puglia.

Figure S2 visualizes the number of observations assigned to the train and test sets, based on the original dataset – i.e., with no synthetic upsampling techniques applied — across all years. As pointed out in the Methods section, we have applied a stratified train-test assignment in order to maintain the same proportion of positive and negative targets in both sets. This also resulted in yearly train-test ratios that are approximately constant, as shown in the graph.

Figure S3 displays the count of infiltrated municipalities (i.e., those with Y=1) in the train and test sets, along with the total. The observable trends show that the number of infiltrated municipalities grew over time consistently up to 2017, and then it declines.

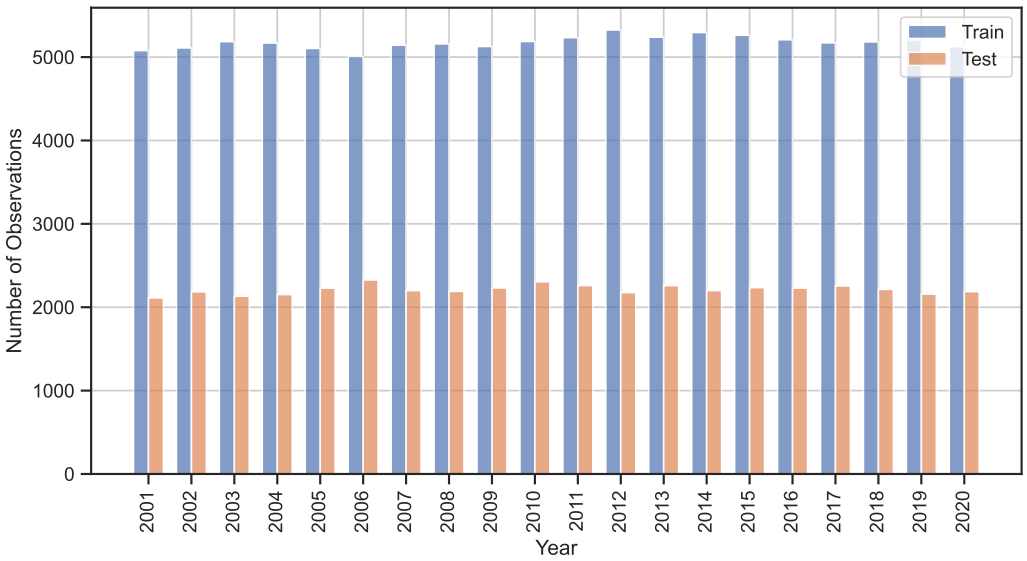
Figure S4 is divided into two subfigures which provide further detail on the distribution of infiltration across regions in Italy, from two different perspectives. On the left we observe, for each region, the percentage of observations (i.e., municipality-year) that are labeled as infiltrated by mafias. In the right subfigure we describe how infiltrated municipalities are distributed across regions.

FIGURE S1. Dismissal trends (overall and by region)



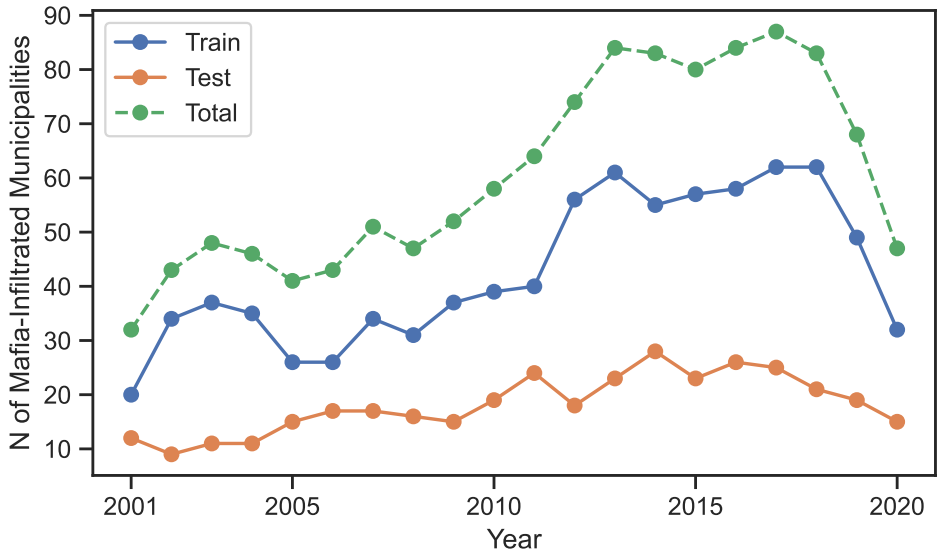
Note: The two figures count the number of actual dismissal decrees issued each year. Subplot (a) shows the overall trend, subplot (b) shows the breakdown by region, including regions that had none over the 2001-2020 period.

FIGURE S2. Distribution of observations in the train and test sets, per each year



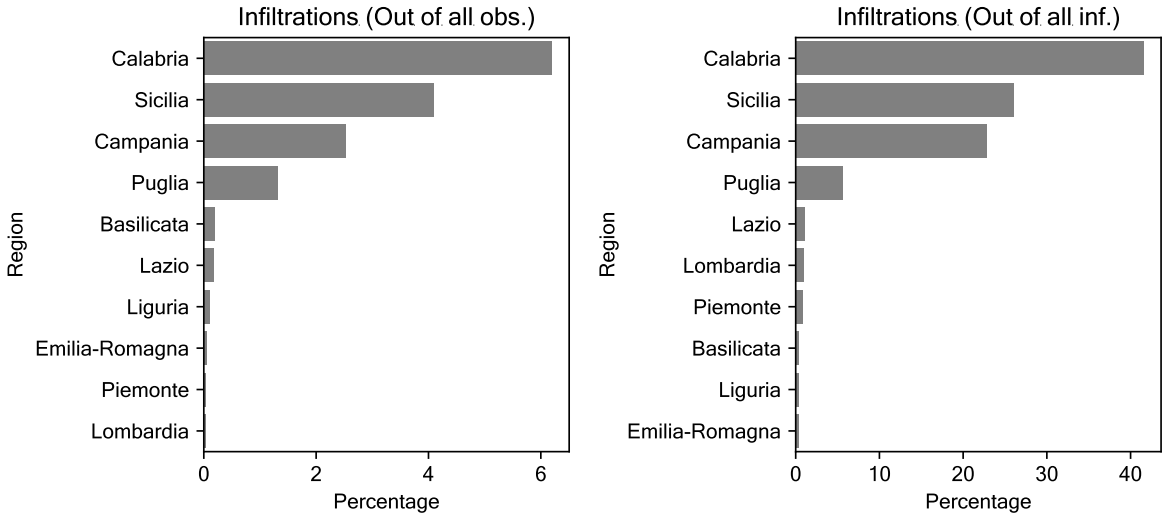
Note: The graph refers to the original dataset (no artificial upsampling approach applied) and shows the count of observations in train and test per each year. The distribution indicates that, as an indirect byproduct of the stratified sampling over our target variable, we also maintain yearly proportions.

FIGURE S3. Number of infiltrated municipalities in the train set, test set and in overall sample.



Note: The figure shows the trend in the number of infiltrated (Y=1) municipality-year observations for the train (blue) set, test (orange) set as well as the entire sample (green). The trends for the train and test sets mimic the underlying trend of the entire sample, which displays an overall increasing number of infiltrated municipalities until 2017.

FIGURE S4. Region-wise percentage of infiltrated municipalities out of all observations (left) and out infiltrated observations only (right).



Note: The figure on the left shows the percentage of infiltrated municipality-year observations over the total number of observations per each region. Only regions that have at least one infiltrated municipality are displayed. The figure on the right shows the region-wise percentage of municipality-year observations out of all infiltrated observations.

A.2 Reducing Class Imbalance

SMOTE. The Synthetic Minority Over-sampling Technique (SMOTE), is a popular method used to address the issue of class imbalance in machine learning datasets (Chawla et al., 2002). Class imbalance in our case refers to the highly unbalanced distribution between negative (i.e., non-infiltrated) and positive (i.e., infiltrated) municipalities. SMOTE works by generating synthetic samples for the minority class to balance the class distribution. This technique, as the other two compared in this work, has been applied only to the training set.

Let D be the dataset with features $\mathbf{X}$ and corresponding class labels y, where $y_i = 1$ indicates the minority class. SMOTE operates by selecting a minority class instance $\mathbf{x_i}$ and its k-nearest neighbors in feature space. A synthetic instance $\mathbf{x_{new}}$ is then created by interpolating between $\mathbf{x_i}$ and one of its neighbors, with the ratio controlled by a parameter $0 \leq \lambda \leq 1$. The synthetic instance is given by:

$$\mathbf{x_{new}} = \mathbf{x_i} + \lambda \times (\mathbf{x_{nn}} - \mathbf{x_i})$$

where $\mathbf{x_{nn}}$ is a randomly selected neighbor of $\mathbf{x_i}$. This process is repeated until the desired balance between classes is achieved.

ADASYN. Adaptive Synthetic Sampling (ADASYN) is an additional technique used to tackle class imbalance in machine learning datasets (Haibo He et al., 2008). Like SMOTE, ADASYN aims to generate synthetic samples, expanding it by introducing adaptability to the density of synthetic samples based on the local distribution of the minority class.

ADASYN identifies the minority instances with fewer synthetic samples and focuses on generating more synthetic samples for those instances. The synthesis process is similar to SMOTE, involving the selection of a minority instance $\mathbf{x_i}$ and its k-nearest neighbors.

However, in ADASYN, a density distribution is computed for each instance, indicating the relative imbalance in its local neighborhood. More synthetic samples are generated for instances in regions with higher density, adapting to the varying degrees of minority class imbalance.

ADASYN aims to overcome some limitations of SMOTE by adapting to the local data distribution, providing a more fine-grained approach to addressing class imbalance.

SMOTE-Tomek. This specific technique refers to the combination of SMOTE (Synthetic Minority Over-sampling Technique) with Tomek links, aiming to improve the quality of synthetic samples generated by SMOTE (Batista et al., 2004).

SMOTE-Tomek first applies SMOTE to oversample the minority class. This involves selecting a minority class instance $\mathbf{x_i}$ and its k-nearest neighbors in feature space.

After applying SMOTE, Tomek links are used to identify pairs of instances (one from the minority class and one from the majority class) that are close to each other. These pairs are considered Tomek links and are used to identify and remove noisy and borderline instances.

SMOTE-Tomek thus combines the over-sampling capability of SMOTE with the cleaning power of Tomek links, aiming to create a more balanced and robust dataset for training machine learning models.

Table S1 presents the distribution of negative and positive observations in the original training test along with the resulting balanced distributions after the application of SMOTE, ADASYN and SMOTE-Tomek.

TABLE S1. Distribution of non-infiltrated (Y=0) and infiltrated (Y=1) city councils in the training set, across sampling approaches

Sampling	Y=0	Y=1	% (Y=1)
Original	105,628	855	0.80
SMOTE	105,628	105,628	50.00
ADASYN	105,628	105,697	50.01
SMOTE+Tomek	105,628	105,628	50.00

Note: The table reports the number of non-infiltrated (Y=0) and infiltrated (Y=1) observations in the train sets for the original dataset as well as the three ones generated using SMOTE, ADASYN and SMOTE+Tomek as artificial upsampling techniques. The rightmost column reports the percentage of infiltrated observations over the total number of observations in each training set.

A.3 Tested Algorithms and Hyperparameter Grid Search

In our predictive framework, we have optimized recall across nine different algorithms, grid-searching for relevant hyperparameters resulting in a total of 50 tested architectures, multiplied by four data sampling techniques, amounting to a total of 200 models. Below, we briefly describe

the nine different algorithms and Table S2 reports the sets of hyperparameters searched across models.

Logistic Regression. Logistic regression is often regarded as the baseline approach in machine learning classification tasks, deployed to test whether a specific task can be meaningfully addressed without resorting to excessively complex models. In our work, we have considered two types of hyperparameters: the penalty type, which we set to 0, and the solver, testing both Saga, which performs better with larger datasets, and lbfgs, which can be more advantageous for small ones.

Ridge Regression. Ridge regression is a regularization technique also known as $L2$- Regularization that gained diffused popularity in machine learning (Hoerl and Kennard, 2000). In Ridge regression, the overarching goal is to stabilize the model by introducing a penalty term during the optimization process. The Ridge penalty encourages the model to identify optimal coefficients that not only minimize the error between predicted and actual values but also penalize the magnitude of the coefficients using the squared $L2$-norm. In our predictive framework, we tested two penalty terms for Lasso Regression, namely 1 and 5, which amounted to λ values equal to 1 and 0.2.

Lasso Regression. Lasso (Least Absolute Shrinkage and Selection Operator) regression is another shrinkage method that uses regularization to address some of the issues that may emerge when using logistic regression for classification tasks (Tibshirani, 2011). The Lasso regularization is also known as $L1$-Regularization, encouraging simpler models with fewer parameters. Compared to Ridge regression (and $L2$-Regularization), when the penalty is sufficiently high, coefficients are set to 0. For this property, LASSO is often used as a feature selection technique before the actual modeling. In the grid search phase, we tested two penalty terms for Ridge, namely 1 and 5, equaling respectively to λ values of 1 and 0.2.

Elastic Net. Elastic Net regression encapsulates the advantages of both Ridge and Lasso regressions (Zou and Hastie, 2005). It combines the $L1$-Regularization from Lasso and the $L2$-

Regularization from Ridge, providing a balanced approach to address the limitations associated with linear regression models. During the grid search phase, we used the elastic net penalty with two penalty terms (i.e., 1 and 5, equaling to λ values of 1 and 0.2) and two L1 ratio parameters, mapping the Elastic-Net mixing parameter (α), being equal to 0.2 and 0.5.[54]

Neural Network. We also tested a multi-layer perceptron (MLP). The MLP is a feedforward artificial neural network consisting of fully connected layers of neurons that are activated by nonlinear functions (Bishop, 1995). Substantially, an MLP is an extension of simple logistic regression, given that the former can stack, in sequence, multiple hidden layers. In the MLP, the input layer consists of input features, i.e., the features in our feature space, while the output layer simply obtains values from the last hidden layer and maps them into prediction (or output) values. In the present work, we tested a simple MLP with two hidden layers, with sizes 6 and 7, and two different numbers of iterations, namely 200 and 500.

Decision Trees. Decision trees are a nonparametric approach designed to infer decisions based on a hierarchical tree-based architecture (Quinlan, 1986). Here, we have searched across a range of five relevant hyperparameters, leading to a pool of 72 potential model candidates. Specifically, we tested maximum depths[55] equal to 5,10, 20; minimum sample split[56] of 5, 10 and 20, minimum sample leaf[57] of 2 and 5. Furthermore, we assessed performance through two split criteria (Gini impurity[58] and Entropy via Shannon's information gain)[59] and two ways to determine the maximum number of features to evaluate when considering the best split (logarithmic and the squared root of the total number of features).

[54]When the L1 ratio is equal to 0, the model becomes automatically a Ridge regression, while when it is set to 1, the model is equal to Lasso regression.

[55]Maximum depth quantifies the the length of the longest path from the root node to a leaf node.

[56]This hyperparameter indicates the minimum number of observations needed to split an internal node

[57]This hyperparameter represents the minimum number of observations that are needed to be at a given leaf node.

[58]Gini impurity in decision trees measures the probability of misclassifying a randomly chosen element's label within a node. It quantifies the impurity or disorder in the node, with lower values indicating a purer (more homogeneous) node. When evaluating a split, the decrease in Gini impurity signifies the improvement in classification, guiding the algorithm to select features that lead to more coherent and accurate decisions.

[59]Information gain quantifies the information gained, with higher values indicating a more informative split. By selecting features that maximize information gain, the algorithm prioritizes those that contribute the most to refining predictions, enhancing the decision-making process.

Random Forests. Random forests is an ensemble method or a meta estimator given that the algorithms is made up of multiple decision trees (Breiman, 2001). During training, random forests build multiple decision trees on random subsets of both the training data and features. This introduces diversity among the trees and helps prevent overfitting to specific patterns in the data. The final prediction is then obtained by aggregating the predictions of these individual trees. Here, we assessed random forests' performance by searching across six relevant hyperparameters: first, the number of estimators (i.e., the number of trees) being either 50 or 100. Then, similar to what has been done with the decision trees, we have searched through minimum samples split (i.e., 5,10,20), minimum sample leaf (i.e., 2,5), split criteria (i.e., Gini and Entropy), and maximum number of features (logarithmic or square root of total features).

Gradient Boosting Machines. Gradient Boosting Machines (GBM) is an ensemble method that has gained wide popularity in the last decade (Friedman, 2001). GBM are part of the boosting family: boosting revolves around the idea of fitting weak learners sequentially, such that at each iteration a new model is trained considering the error of the previously fitted model. In this regard, GBM differ from traditional ensemble methods such as Random Forests, which generates final predictions by simply averaging the performance of each single tree. In the present work, we sampled from a pool of 24 possible GBM as the result of a hyperparameter search across four different hyperparameters: number of estimators (i.e., 50 or 100), maximum depth (i.e., 5, 10, 20), maximum number of features to be selected at each iteration (selected via logarithmic or square root sampling) and learning rate (i.e., 0.1 or 0.2).

XGBoost. Extreme Gradient Boosting (XGBoost) (Chen and Guestrin, 2016) is also defined as a tree-based ensemble method and, similarly to GBM, is also built on the principles of boosting such that each weak learner is trained taking into considerations the errors of previous weak learners. Differently from GBM, however, XGBoost relies on a scalable and highly efficient implementation that reaches considerably higher computational speed. In fact, while GBM trains weak learners sequentially, XGBoost does so in parallel. Additionally, XGBoost incorporates regularization techniques to control the complexity of the individual weak learners and prevent overfitting. An overall population of 144 potential model candidates has been considered for

the grid search of five relevant hyperparameters. These were: number of estimators (i.e., 50, 100), maximum tree depth (i.e., 5, 10, 20), weight to positive sample (i.e., 1, 10, 50), learning rate (i.e., 0.1, 0.2) as well as α (i.e., 0, 0.5)[60] and γ (i.e., 0, 0.5).[61]

[60] α maps the L1 regularization term.

[61] γ maps the minimum loss reduction that is required to add a partition on a leaf node of a given tree.

TABLE S2. List of tested algorithmic methods along with compared hyperparameters spaces and total number of candidate models for the potential pool.

Algorithm	**Hyperparameters**	**N of Model Candidates**
Logistic Regression	Penalty type=[no penalty], Solver=[Saga, lbfgs]	2
Ridge Regression	Penalty type=[L2], Penalty terms=[1,5], Solver=[Saga, lbfgs]	4
Lasso Regression	Penalty type=[L1], Penalty term=[1,5], Solver=[Saga]	2
Elastic Net	Penalty type=[elastic net], Penalty terms=[1,5], L1 ratio=[0.2, 0.5], Solver=[Saga]	2
Neural Network	N of hidden layers=[2], Hidden layers sizes=[(6,7)], Activation=[logistic], Solver=[Adam], Max N of iteration=[200,500]	2
Decision Tree	Max depth=[5,10,20], Min samples split=[5,10,20], Min samples leaf=[2,5], Split criterion=[Gini, Entropy], Max N of features=[log2, automatic]	72
Random Forest	N of estimators=[50,100], Max depth=[5,10,20], Min samples split=[5,10,20], Min samples leaf=[2,5], Split criterion=[Gini, Entropy], Max N of features=[log2, automatic]	144
Gradient Boosting Machines	N of estimators=[50,100], Max depth=[5,10,20], Max N of features=[log2, automatic], Learning rate=[0.1, 0.2]	24
XGBoost	N of estimators=[50,100], Max depth=[5,10, 20], Weight to positive samples=[1,10,50], Learning rate=[0.1, 0.2], Alpha=[0, 0.5], Gamma=[0, 0.5]	144

Note: This table reports information regarding the various utilized algorithms, along with the hyperparameters employed for the optimization phase and the total number of model candidates per each algorithm. Grand total of potential pool model candidates is equal to 394 $\times$ 4 sampling approaches = 1,576. We randomly searched 10 combinations when the number of model candidates (per each algorithm) is higher than 10. In total, the actual number of tested models is thus 200.

A.4 Explainability via SHAP: Technical Details

SHAP (SHapley Additive exPlanations) has gained significant popularity as a tool for interpreting outcomes of machine learning models (Lundberg and Lee, 2017), drawing inspiration from cooperative game theory, specifically Shapley Values introduced by Shapley (1952). The SHAP approach provides a model-agnostic method for explaining predictions, offering nuanced insights into the contribution of each feature to the model's output and enhancing transparency and interpretability.

The Shapley Value for feature i (ϕ_i) is computed by summing over all possible subsets T of the model input space. This formula integrates combinatorial coefficients to ensure a fair distribution of the contribution of feature i across different subsets:

$$\phi_i(S) = \sum_{T \subseteq S} \frac{|T|!(|S| - |T| - 1)!}{|S|!} [f(x_{T \cup \{i\}}) - f(x_T)] \tag{S1}$$

Here, f represents the prediction function of the model, x_T corresponds to the input with features in set T, and $x_{T \cup \{i\}}$ corresponds to the input with features in set T along with feature i. The terms $|T|$ and $|S|$ denote the number of features in set T and the total number of features in the dataset, respectively.

The overall impact of feature i (ϕ_i) is obtained by summing the Shapley Values for feature i across all possible subsets S of features, excluding feature i:

$$\phi_i = \sum_{S \subseteq \{1,2,\ldots,p\} \setminus \{i\}} \frac{|S|!(p - |S| - 1)!}{p!} \phi_i(S) \tag{S2}$$

This formula ensures a comprehensive evaluation of the contribution of feature i by considering its impact across various combinations of other features, where p is the total number of features.

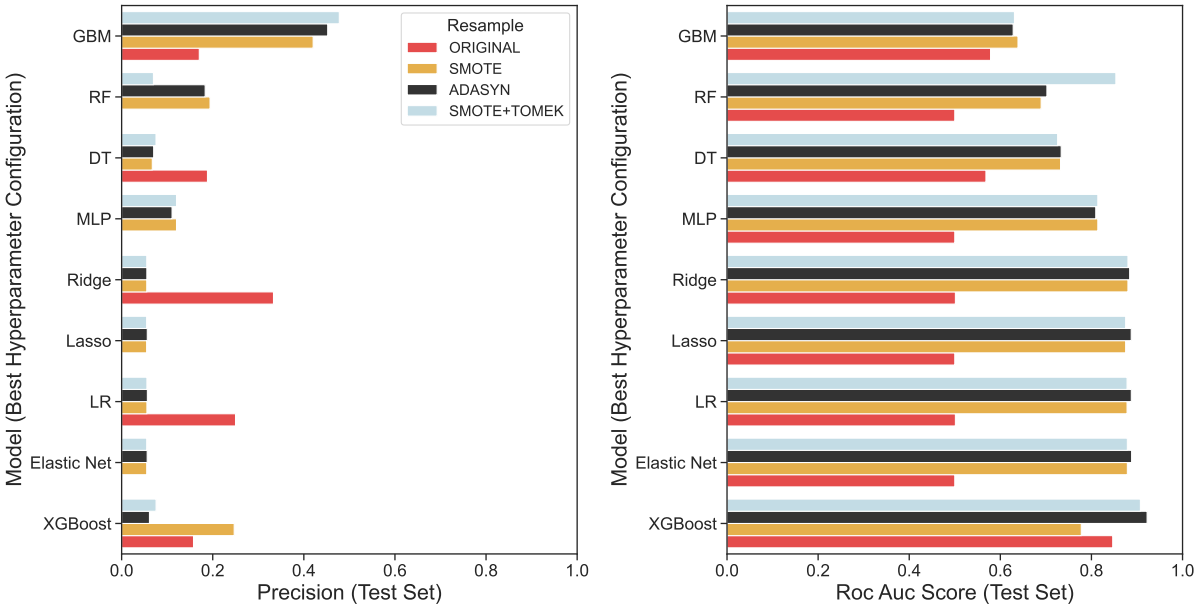
Given the tree-based method achieving the best predictive result, i.e., XGBoost, we specifically used TreeSHAP, an extension of SHAP meant to be more efficient from the computational point of view when dealing with tree-based methods (Lundberg et al., 2019).

A.5 Supplementary Predictive Results

Confusion Matrix. Table S3 reports two confusion matrices obtained after deploying the main model used in the prediction. On the top, we consider all observations in the train and test set, on the bottom we only focus on observations in the test set. The table indicates that, in the test set, only 13 out of 351+13 truly infiltrated municipality-year observations have not been predicted as such, amounting to a recall of 0.96. In both cases, around 12% of observations are predicted as infiltrated out of the total (12.58% in the whole sample, 12.84% in the test set). This further details how our machine learning algorithm provides a high number of false positives, when focusing on the 0.5 threshold.

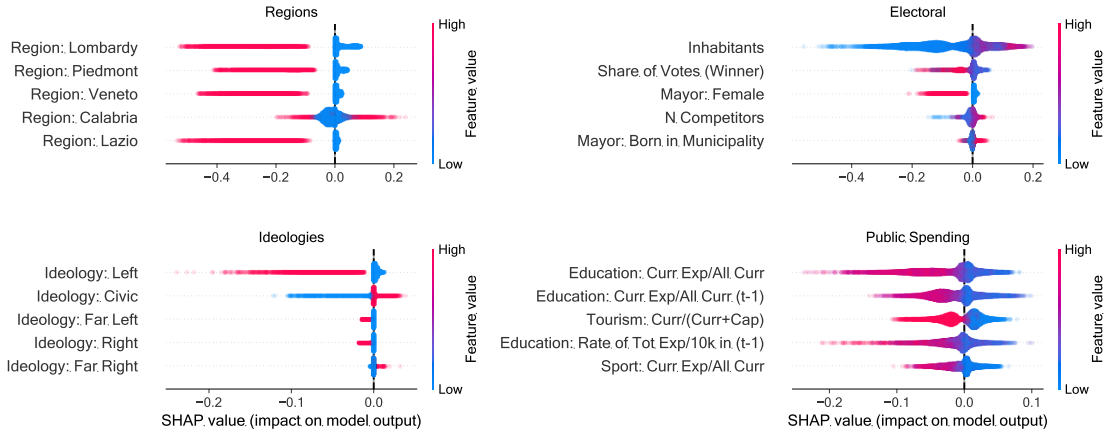
SHAP Wrongful Dismissals. In the paper, we discuss our ability to predict wrongful dismissals. Here, we also show (Figure S10) the ten most important features according to SHAP. Two regional variables emerge as relevant, Sicily and Calabria, positively correlated with the probability of a correct dismissal. Having more candidates and an incumbent mayor are also associated with correct dismissals. We then observe some public spending variables. A notable one is the share of current expenditures on culture over total spending on this item. This variable is negatively associated with correct dismissals. Interestingly, this variable is also associated with a lower risk of dismissal in the main analysis.

FIGURE S5. Predictive performance: Precision and Roc-Auc Scores



Note: Precision and Roc-Auc Scores are computed on the test set, per each model and each sampling approach and refers to the best performing architecture emerged from the randomized grid search in terms of hyper-parameter combination. Models are ordered based on increasing Roc-Auc Score for the ADASYN resample. XGBoost with ADASYN sampling reaches the highest Roc-Auc performance (0.92).

FIGURE S6. SHAP summary plot - best performing machine learning model



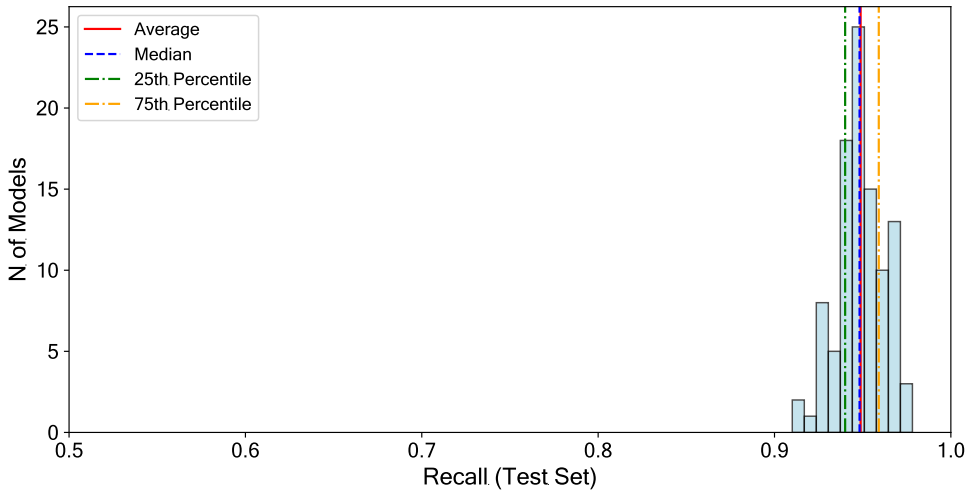
Note: The plot displays the five most important features per each dimension, i.e., Regions, Electoral, Ideologies, Public Spending. The best performing machine learning model is XGBoost with ADASYN as synthetic sampling approach. Each point on the graph corresponds to an observation in the test set, with the color mapping the observation's value relative to the feature's scale. The x-axis quantifies the SHAP value on the probability scale, indicating whether an instance pushes the algorithm towards a positive or negative prediction correlation with mafia infiltration.

TABLE S3. Confusion Matrices – Overall Sample and Test Set Only

Train + Test			
		Predicted	
		Not Infiltrated	Infiltrated
True Label	Not Infiltrated	129,168	17,404
	Infiltrated	13	1,202
Test Set Only			
		Predicted	
		Not Infiltrated	Infiltrated
True Label	Not Infiltrated	38,530	5,329
	Infiltrated	13	351

Note: This Table reports two confusion matrices as obtained by the main analysis. The first one (on the top) uses all observations in the train and test sets. The second (on the bottom) instead only uses observations in the test set.

FIGURE S7. Distribution of Recall Scores (on the Test Set) over Sensitivity Models



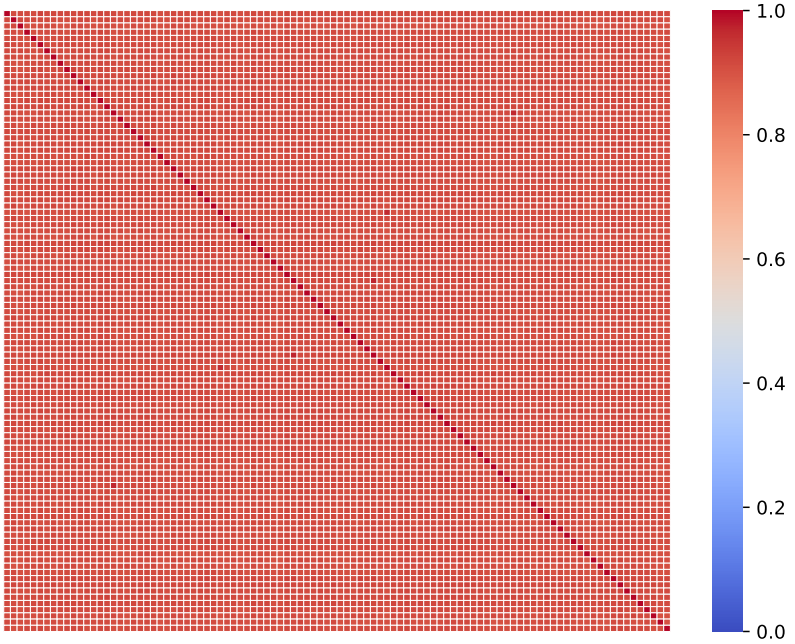
Note: This plot shows the distribution in terms of recall on the test set for the 100 sensitivity models run with different random seed combinations used for the train and test split phase and the artificial upsampling phase via ADASYN.

TABLE S4. Out-of-sample (i.e., on test set) performance of XGBoost with best model hyperparameter combination with stratified train-test split (i.e., without temporal structure)

Feature Space	Test Set Size	Recall	Roc-Auc	Precision
All Features	0.1	0.991	0.929	0.056
All Features	0.2	0.963	0.918	0.057
All Features	0.3	0.964	0.922	0.061
No Regions	0.1	0.959	0.884	0.039
No Regions	0.2	0.913	0.867	0.039
No Regions	0.3	0.896	0.861	0.040
No Electoral	0.1	0.983	0.920	0.053
No Electoral	0.2	0.946	0.905	0.053
No Electoral	0.3	0.948	0.914	0.060
No Ideology	0.1	0.983	0.928	0.059
No Ideology	0.2	0.967	0.920	0.058
No Ideology	0.3	0.983	0.928	0.059
No Public Spending	0.1	1.000	0.729	0.014
No Public Spending	0.2	1.000	0.717	0.014
No Public Spending	0.3	0.997	0.727	0.014

Note: This table reports the results of different predictive exercises using XGBoost with the same hyperaparameter combination as the one used for the main analyses. We compared the predictive performance of fifteen models in terms of recall, Roc-Auc score and precision when keeping all feature as well as removing different sets of features and testing different test set sizes. Specifically, we first used all features, and then removed a) regional dummies, b) electoral variables, c) ideology variables, and d) public spending variables. In terms of test set sizes, the models use respectively a) 10% of observations, b) 20% of observations and c) 30% of observations. The model using all features and a test set size of 30% is the main model commented in the main text.

FIGURE S8. Correlation of Predicted Probabilities over Sensitivity Models



Note: This plot shows the Pearson correlation of the predicted probabilities for the 100 sensitivity models run with different random seed combinations used for the train and test split and the artificial upsampling phase via ADASYN.

TABLE S5. Out-of-sample (i.e., on test set) performance of XGBoost with best model hyperparameter combination with non-stratified train-test split (i.e., with temporal structure)

Feature Space	Test Set Size	Recall	Roc-Auc	Precision
All Features	0.101 (from 2019 on)	0.975	0.923	0.057
All Features	0.203 (from 2017 on)	0.935	0.906	0.058
All Features	0.304 (from 2015 on)	0.946	0.912	0.059
No Regions	0.101 (from 2019 on)	0.943	0.878	0.039
No Regions	0.203 (from 2017 on)	0.907	0.866	0.040
No Regions	0.304 (from 2015 on)	0.903	0.866	0.041
No Electoral	0.101 (from 2019 on)	0.951	0.902	0.050
No Electoral	0.203 (from 2017 on)	0.943	0.903	0.053
No Electoral	0.304 (from 2015 on)	0.935	0.898	0.051
No Ideology	0.101 (from 2019 on)	0.967	0.919	0.057
No Ideology	0.203 (from 2017 on)	0.955	0.916	0.059
No Ideology	0.304 (from 2015 on)	0.948	0.913	0.059
No Public Spending	0.101 (from 2019 on)	0.983	0.735	0.015
No Public Spending	0.203 (from 2017 on)	0.987	0.720	0.014
No Public Spending	0.304 (from 2015 on)	0.983	0.713	0.014

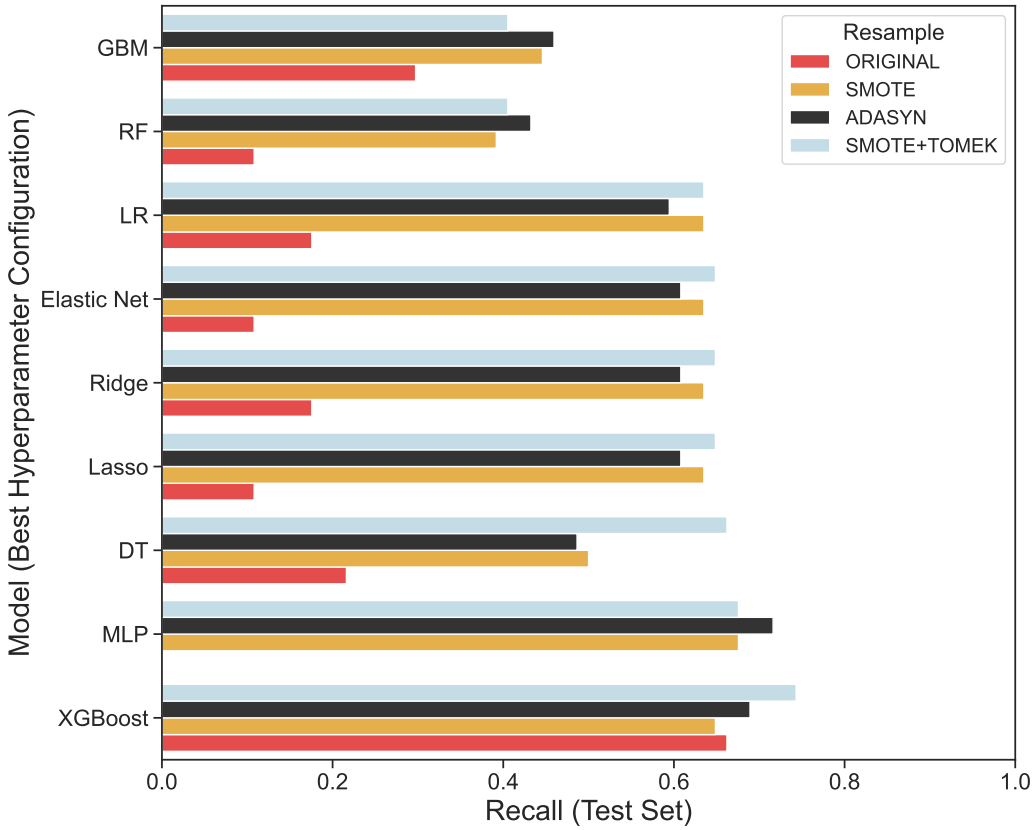
Note: This table reports the results of different predictive exercises using XGBoost with the same hyperaparameter combination as the one used for the main analyses. We compared the predictive performance of fifteen models in terms of recall, Roc-Auc score and precision when removing different sets of features and testing different test set sizes maintaining the temporal order of the observations. Specifically, we first used all features, and then removed a) regional dummies, b) electoral variables, c) ideology variables, and d) public spending variables. In terms of test set sizes, the models use respectively a) data from 2019 on in the test set (amounting to 10.1% of observations), b) data from 2017 on (amounting to 20.3% of the observations) and c) data from 2015 on (amounting to 30.4% of the observations).

TABLE S6. 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
Municipality Dummies	152,119 (0.80)	Original	0.776	0.865	0.120
		ADASYN	0.934	0.901	0.054
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 dismissal)	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

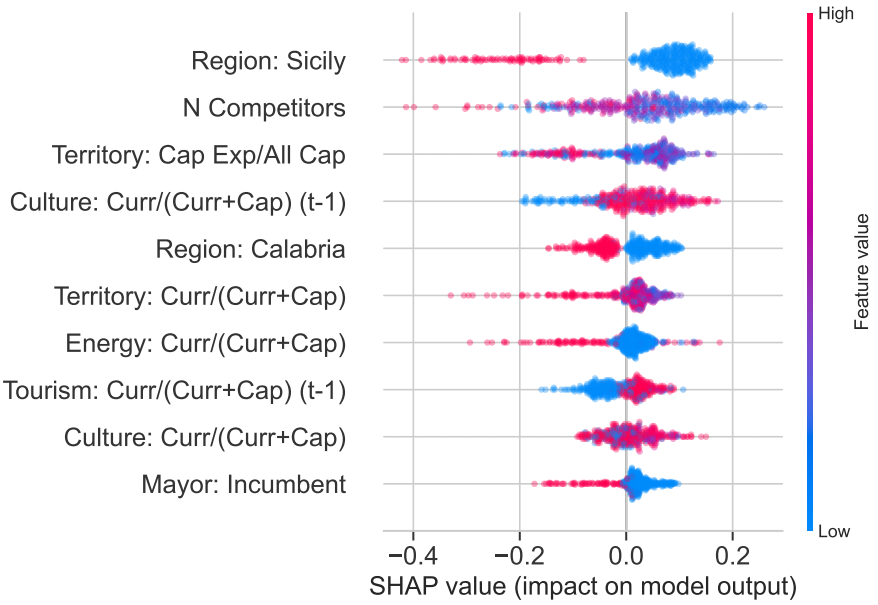
Note: This table reports results for five different predictive exercises. All models use XGBoost with the hyper-parameter configuration of the best model deployed in the main analyses. For symmetry all exercises have been performed using both the original sampling as well as ADASYN. The first configuration (see leftmost column) only includes municipalities from southern regions which are traditionally those more plagued by mafia as well as those where detection in higher. The second includes binary municipality-level features instead of regional ones. The third configuration only retains municipalities that have been dissolved once. The fourth configuration retains municipalities that have been dissolved once and removes all observations (years) after infiltration has been detected. For this fourth configuration, we removed 12 municipalities that have been dissolved in two non-consecutive terms. The fifth configuration considers a cross-sectional sample, partially following the approach of Colonnelli et al. (2020). Per each model, we report recall, Roc-Auc score and precision. All metrics refer to the test set.

FIGURE S9. Predictive performance (Wrongful vs Correct dismissals): Recall Scores



Recall scores are computed on the test set, per each model and each sampling approach and refers to the best performing architecture emerged from the randomized grid search in terms of hyperparameter combination. Models are ordered in terms of increasing Recall for SMOTE-Tomek resampling. XGBoost with SMOTE-Tomek sampling reaches the highest Roc-Auc performance (0.74)

FIGURE S10. SHAP summary plot - best performing machine learning model (task: predicting wrongful vs correct dismissals).



Note: the plot displays the ten most important features overall. The best performing machine learning model is XGBoost with SMOTE-Tomek as synthetic sampling approach. Each point on the graph corresponds to an observation in the test set, with the color mapping the observation's value relative to the feature's scale. The x-axis quantifies the SHAP value on the probability scale, indicating whether an instance pushes the algorithm towards a positive or negative prediction correlation with mafia infiltration.

A.6 A Synthetic Measure of Infiltration Risk – Further Details

Description of the Measure. Figure S11 visualizes the trend in the average predicted infiltration probability, for the train and the test. The picture highlights how the two are highly correlated over time and close to each other. It also shows how using all the observations (from the train and test sets) and associated infiltration risks does not pose the problem of having totally unrelated resulting risks.

While in the main text we report the distribution of the synthetic indicator of mafia infiltration risk for the entire dataset (Figure S12), Figure S13 disaggregates the distribution across four macro-regions: South, Center, North-East, and North-West. South is the macro-region with the higher predicted risk of infiltration. In the South, the mean of the indicator is 0.38, with a median of 0.18. Interestingly, 10% of all municipality-year observations have a predicted risk of 0.99 or higher. The Center macro-region displays a much lower risk of infiltration compared to the South, but nonetheless it appears as the one having the second highest predicted mean (0.07), while North-East and North-West have lower mean risks.

Finally, Figure S14 provides a map of the predicted infiltration at the municipal level for the years 2001, 2010, 2020. The plots demonstrate the sensibly higher number of municipalities at risk in the South.

External Validity: Evidence for Dismissals in 2021–2023. In a further exercise to assess the predictive ability of our machine learning model in correctly predicting municipalities that are detected to be infiltrated by mafias, we have considered all cases of city council dismissals that occurred between 2021 and 2023, and thus in a time range outside of the one used to train and test our algorithms. In principle, this predictive task is more challenging, as we can observe spending patterns only a few years before the dismissal. Table S7 reports the results of our analysis.

A total of 33 municipalities have been dissolved in this time range (14 in 2021, 11 in 2022, 8 in 2023). Using only the 2020 prediction of risk, our model correctly identifies 72% of the municipalities in the entire period (71% of those dissolved in 2021, 63% in 2022, 87% in 2023). If we use the average prediction in the past five years, i.e., 2016-2020, our model correctly

identifies 81% of municipalities dissolved in the entire period (85% in 2021, 72% in 2022, 87% in 2023). Finally, if we take the average risk of the last term, we identify 78% of municipalities that have been dissolved (78% in 2021, 72% in 2022, 87.5% in 2023). In the table, we highlight in red the dismissed municipalities, which we *fail* to predict. Interestingly, when we successfully predict a dismissal, our index reports extremely high scores, often close or equal to 100%.

FIGURE S11. Trend in the average value of the synthetic indicator of mafia risk over the years, for both train and test sets

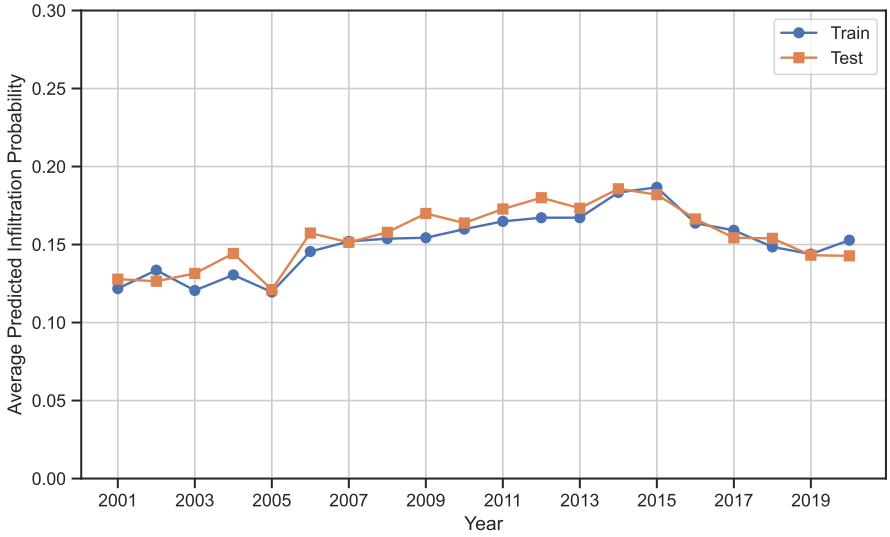
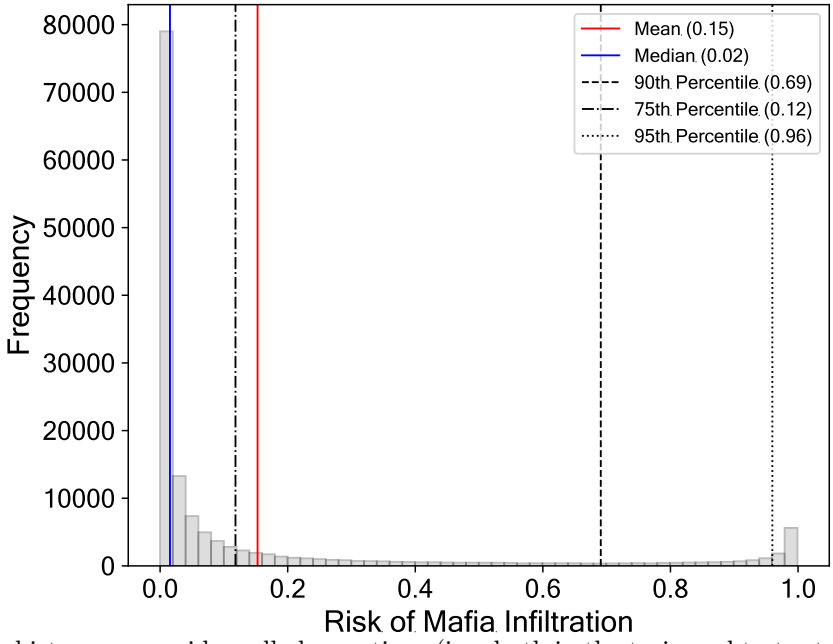
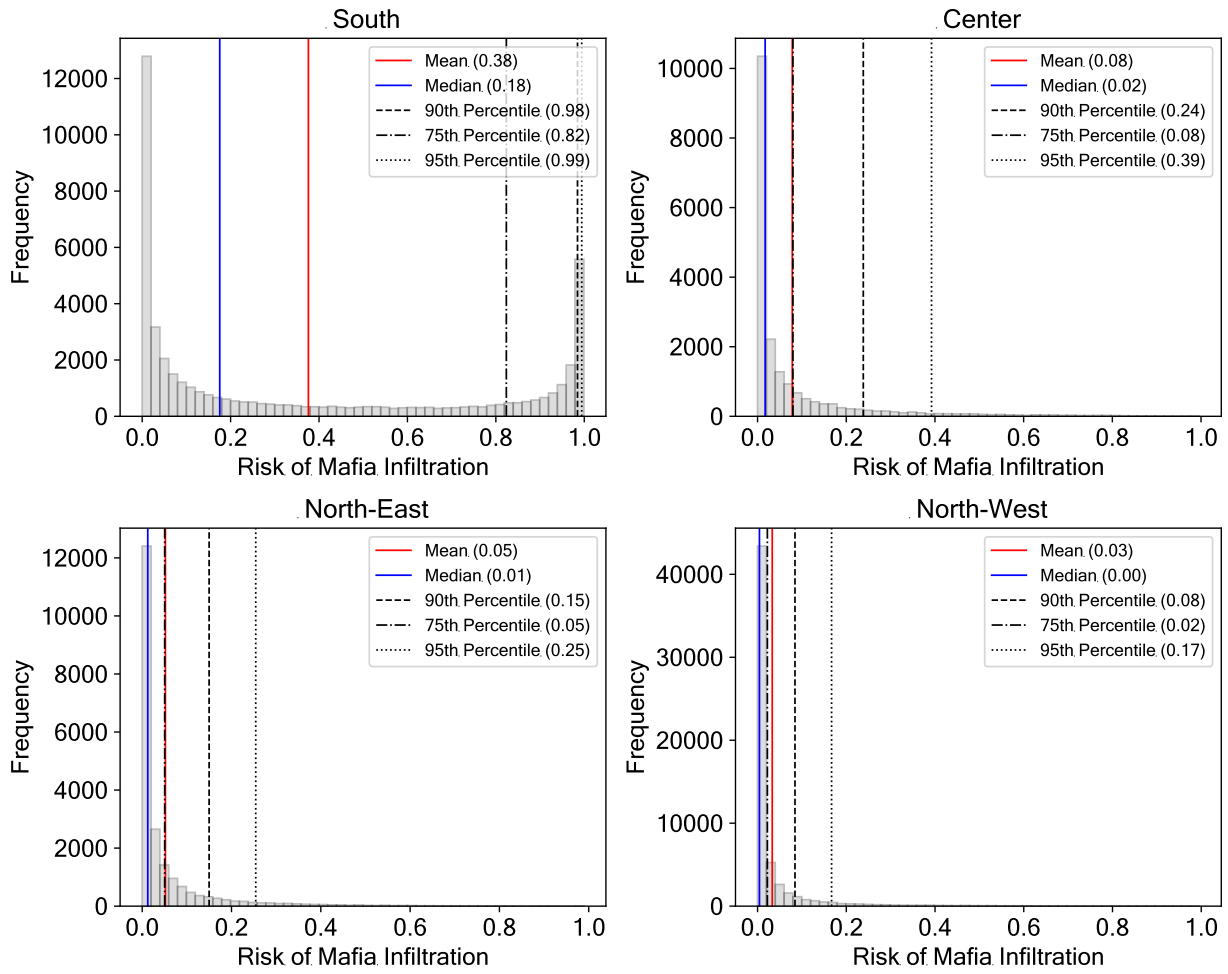


FIGURE S12. Infiltration Risk Distribution (Whole Sample)



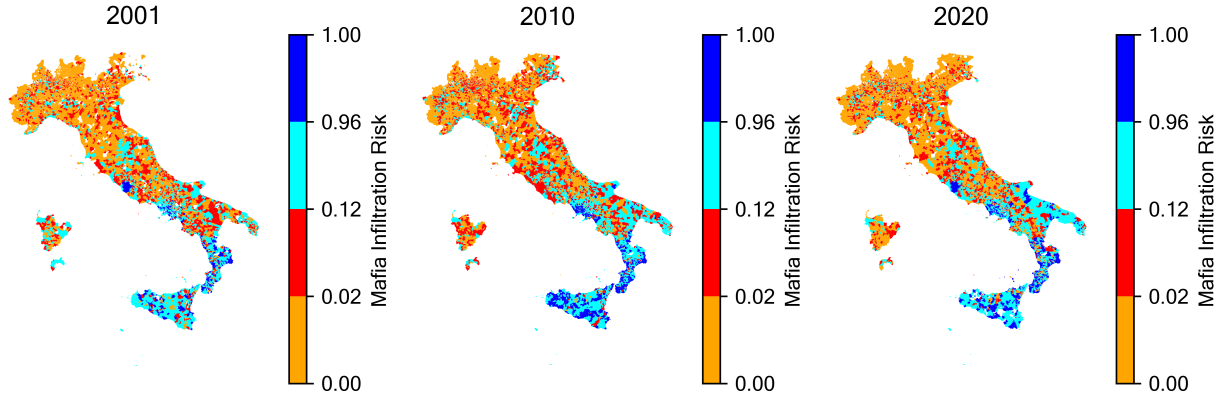
Note: The histogram considers all observations (i.e., both in the train and test sets) in the dataset. Mean and median values are highlighted via vertical lines, along with values related to 75th, 90th and 95th percentiles.

FIGURE S13. Infiltration Risk Distribution (by Macro-Region)



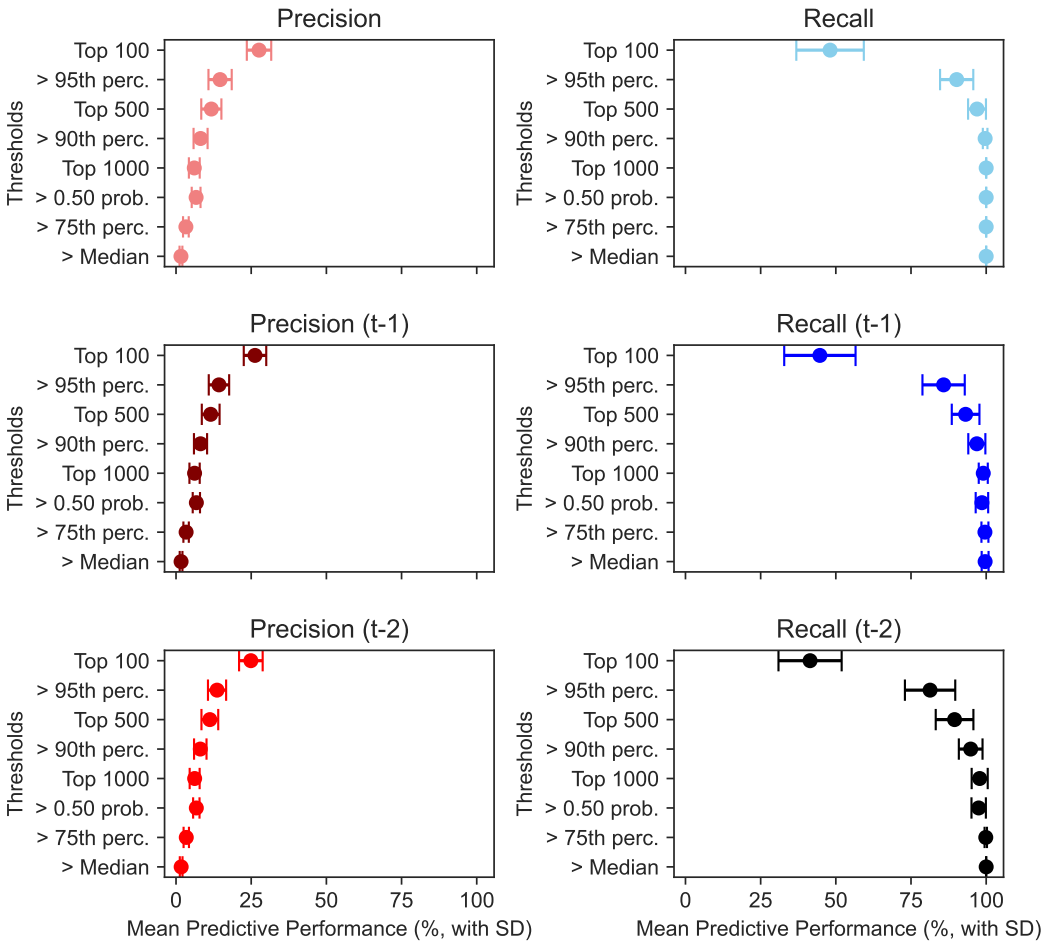
Note: The histogram considers macro-region based distribution of the risk indicator, considering all observations (i.e., both in the train and test sets). Mean and median values are highlighted via vertical lines, along with values related to 75th, 90th and 95th percentiles and refer to each macroregion.

FIGURE S14. Spatial distribution of the risk indicator at the municipal level (2001, 2010, 2020)



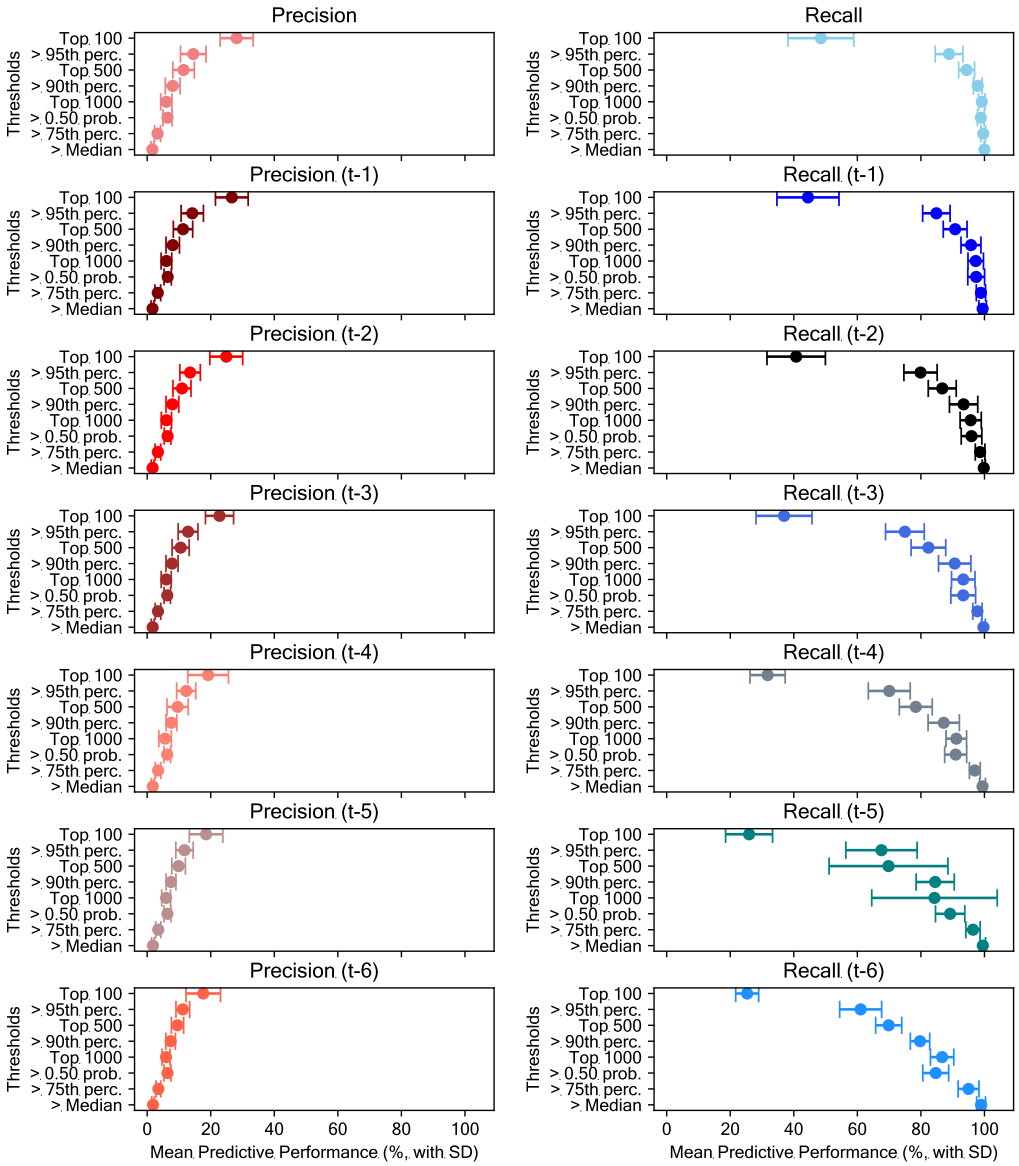
Note: Missing municipalities indicate missing predicted infiltration risk due to lack of electoral data. These observations are mostly present in autonomous regions such as Trentino Alto-Adige, Aosta Valley and Friuli-Venezia Giulia. The colormap has four colours: orange if the predicted risk is between 0 and the overall median value for the 2001-2020 period (i.e., 0.02), red if the risk lies between 0.02 and the value representing the 75th percentile for the entire period (i.e., 0.12), cyan is the risk lies between the 75th and the 95th percentile (i.e., 0.96); blue is the risk is higher than the 95th percentile.

FIGURE S15. Precision and recall for various thresholds and rankings - by year (2001-2020), Risk computed as average after 100 model runs



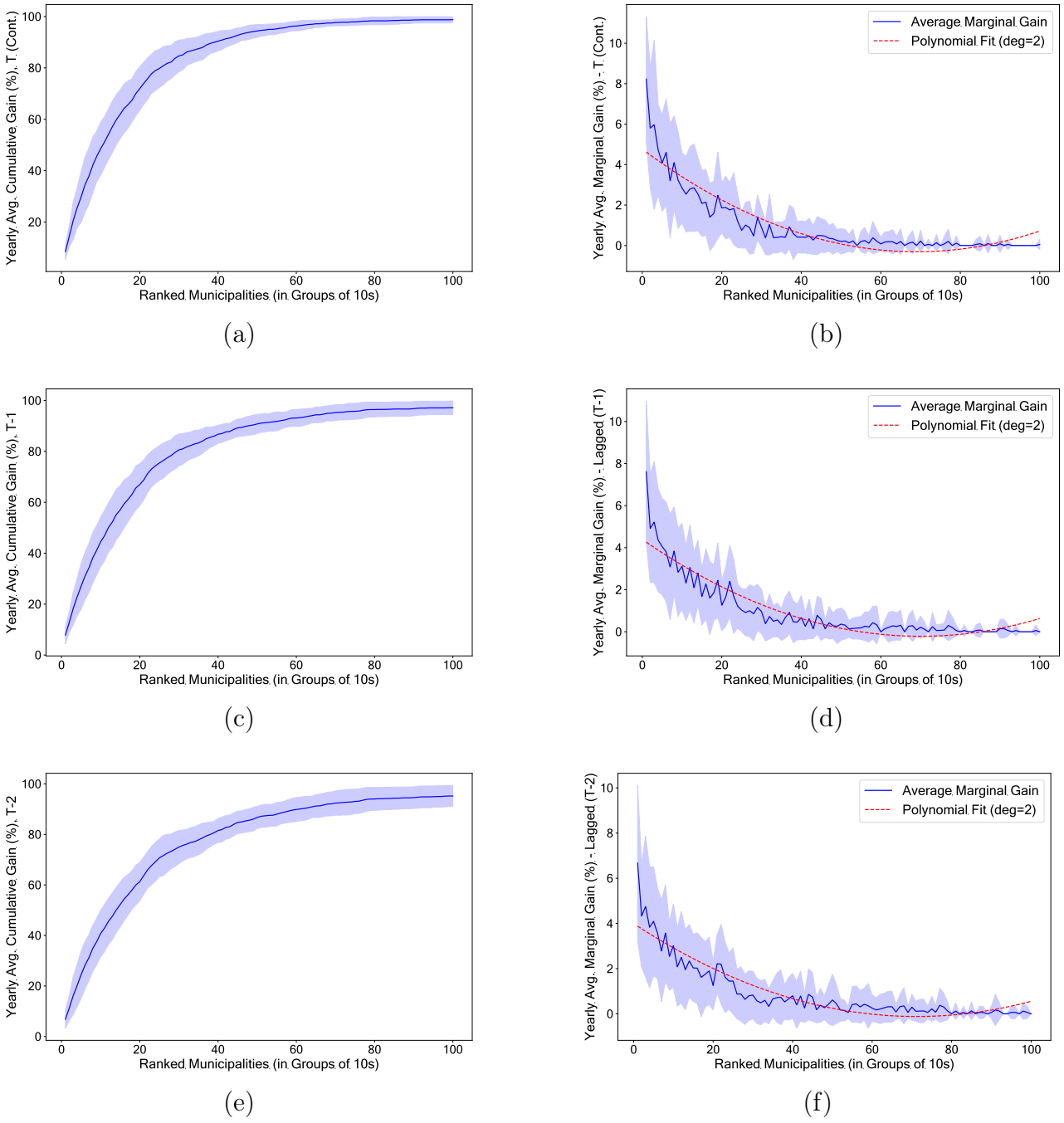
Note: The predicted risk is the average predicted risk obtained after 100 model runs with as many random combinations of random seeds. We report mean and standard deviation. Precision and recall are shown as percentages. Precision (t-1) and Recall (t-1) refer to the performance when using the risk indicator at t - 1 to predict infiltration at t. Precision (t-2) and Recall (t-2) refer to the performance when using the risk indicator at t - 2 to predict infiltration at t. Performance is considered using predicted risk for observations both in the train and test sets. Threshold values are calculated by considering each yearly distribution of the predicted risk of infiltration.

FIGURE S16. Precision and recall for various thresholds and rankings - by year (2001-2020), Risk computed as average after 100 model runs



Note: We report mean and standard deviation. Precision and recall are shown as percentages. Precision (t-n) and Recall (t-n) refer to the performance when using the risk indicator at t - n to predict infiltration at t (e.g., Precision and Recall at t-6 refer to the performance of the index when using its values at t-6 to predict infiltration at t. Performance is considered using predicted risk for observations both in the train and test sets. Threshold values are calculated by considering each yearly distribution of the predicted risk of infiltration.

FIGURE S17. Cumulative and Marginal gain - Recall



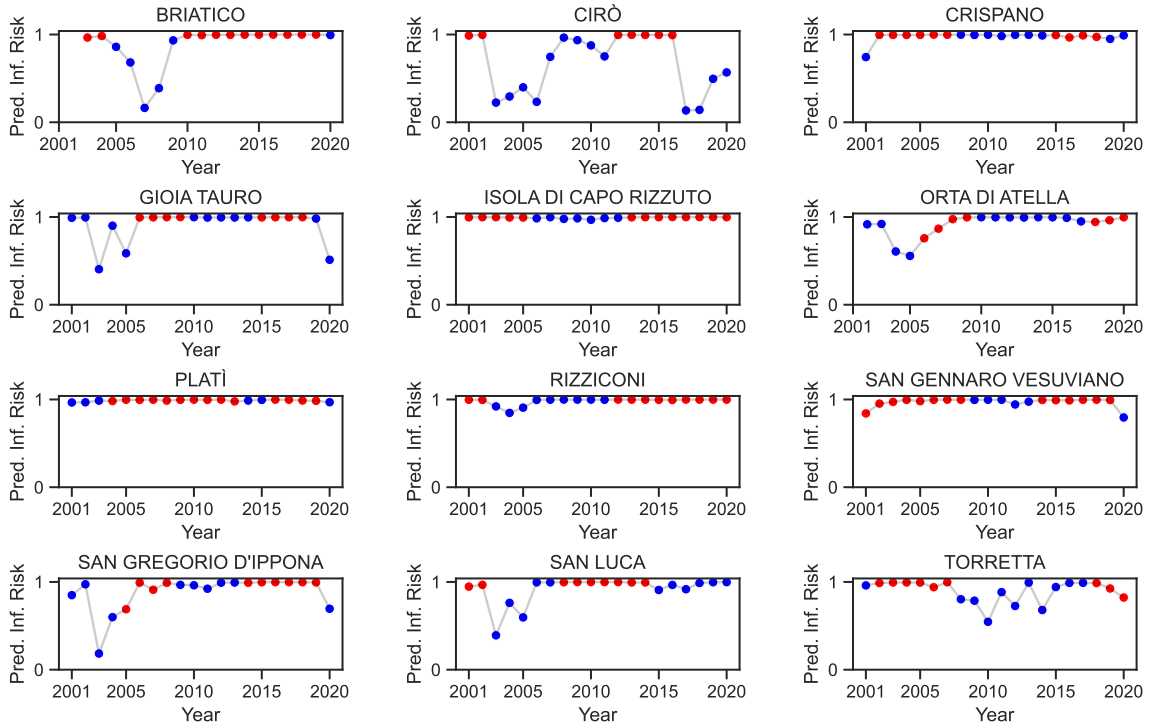
Note: Subplots (a), (c), (e) show the yearly average cumulative gain in terms of recall (displayed in percentage points). Subplots (b), (d), (f) show the marginal gain also in terms of recall. Cumulative and marginal gains were obtained when investigating municipalities in groups of 10s, ranked in decreasing order in terms of predicted risk of infiltration. The top subplot row refers to the cumulative and marginal gains when using the predicted risk at t to measure performance at t. The center subplot row use predicted risk at t-1 to measure performance at t. The bottom subplot row use predicted risk at t-2 to measure performance at t. The shaded areas around the average trend lines of the cumulative gain subplots represent the standard deviation calculated at the yearly level. We censor the ranked municipalities up to the 100th group, i.e., to 1000 municipalities.

TABLE S7. Infiltration risk for municipalities dissolved in 2021-2023.

Municipality	Region	dismissal year	2020	Avg. 2016-2020	Last term
Squinzano	Puglia	30/01/21	98%	92%	97%
Guardavalle	Calabria	23/02/21	100%	100%	100%
Carovigno	Puglia	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	Puglia	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	Puglia	27/12/21	42%	32%	24%
Castellammare di Stabia	Campania	24/02/22	90%	75%	94%
Trinitapoli	Puglia	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	Puglia	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	Puglia	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%

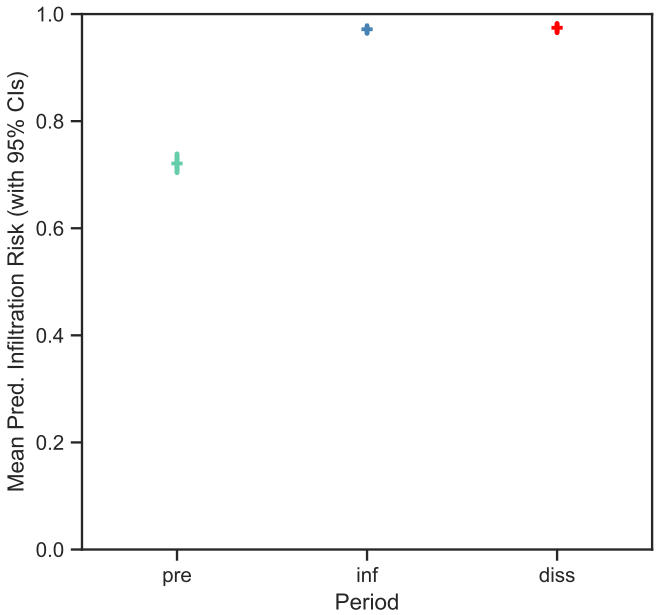
Note: These municipalities were not included in either the training or test phase of our predictive framework. Red cells indicate predicted risk below 50% (hence the algorithm would predict them as non-infiltrated). The indicator is rendered in percentage terms.

FIGURE S18. Predicted Risk of infiltration – Municipalities dissolved more than once in two non-consecutive electoral mandates



Note: The plot only considers those municipalities that have been dissolved more than once and not for two consecutive mandates. Red dots indicate periods in which the municipality was infiltrated, and blue dots indicate periods in which the municipality was not detected as such.

FIGURE S19. Average predicted risk for dissolved municipalities (Pre-infiltration, infiltration, actual dismissal)

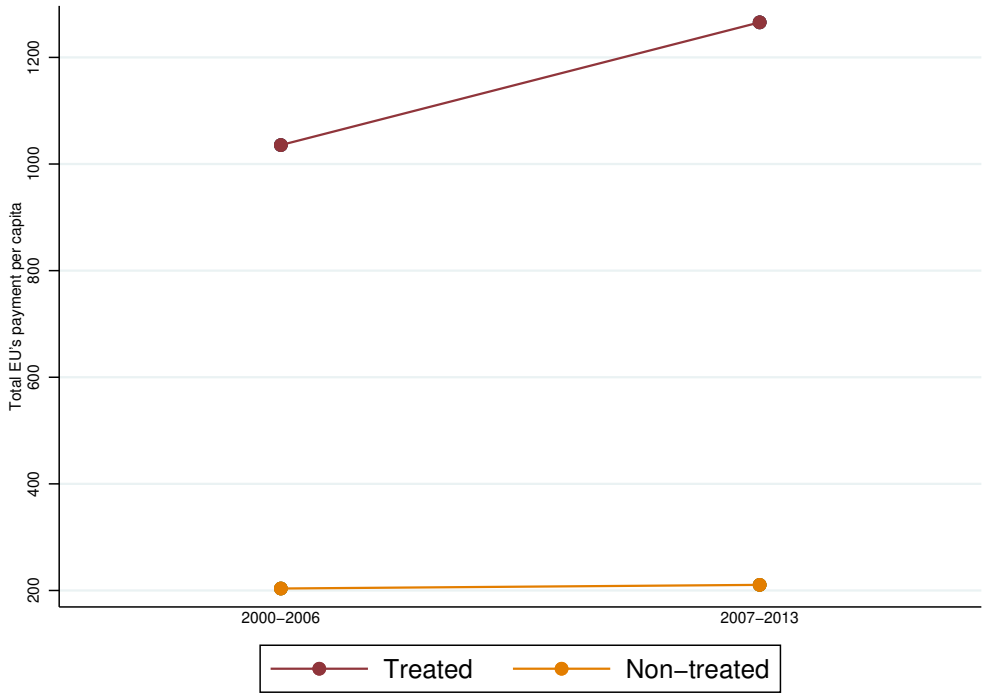


Note: The plot collapses all the pre-infiltration and infiltration years to compare the average distribution for all the municipalities dissolved once (or more than once, but for consecutive electoral mandates) between 2001 and 2020. Diss refers to the mean predicted risk for those municipalities in the years they were actually dissolved.

B EU Funds and Mafia Infiltration

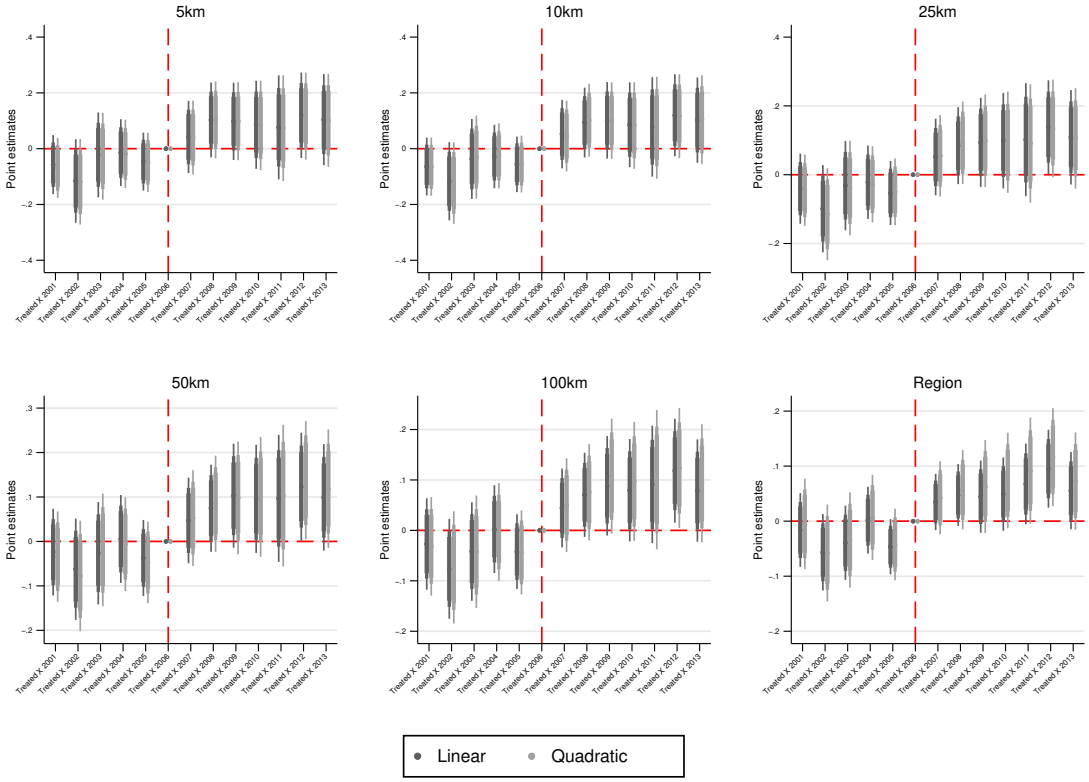
B.1 Additional Figures

FIGURE S20. Amount of EU funds (per capita) between treated and control regions, for different funding periods



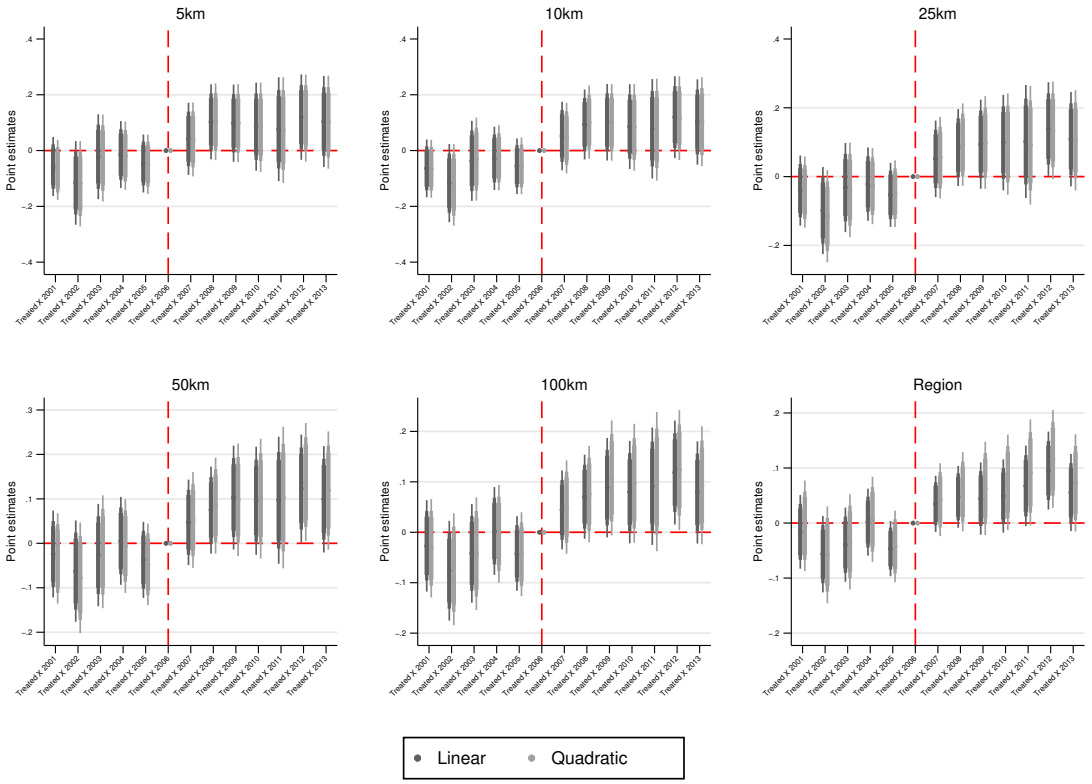
Note: The figure shows the total amount of European funds, in per capita terms, received by the treated and control regions during the funding waves of 2000-2006 and 2007-2013. Data are retrieved from the Euopean Cohesion Data Platform

FIGURE S21. Event study results: main models, with robust coefficients



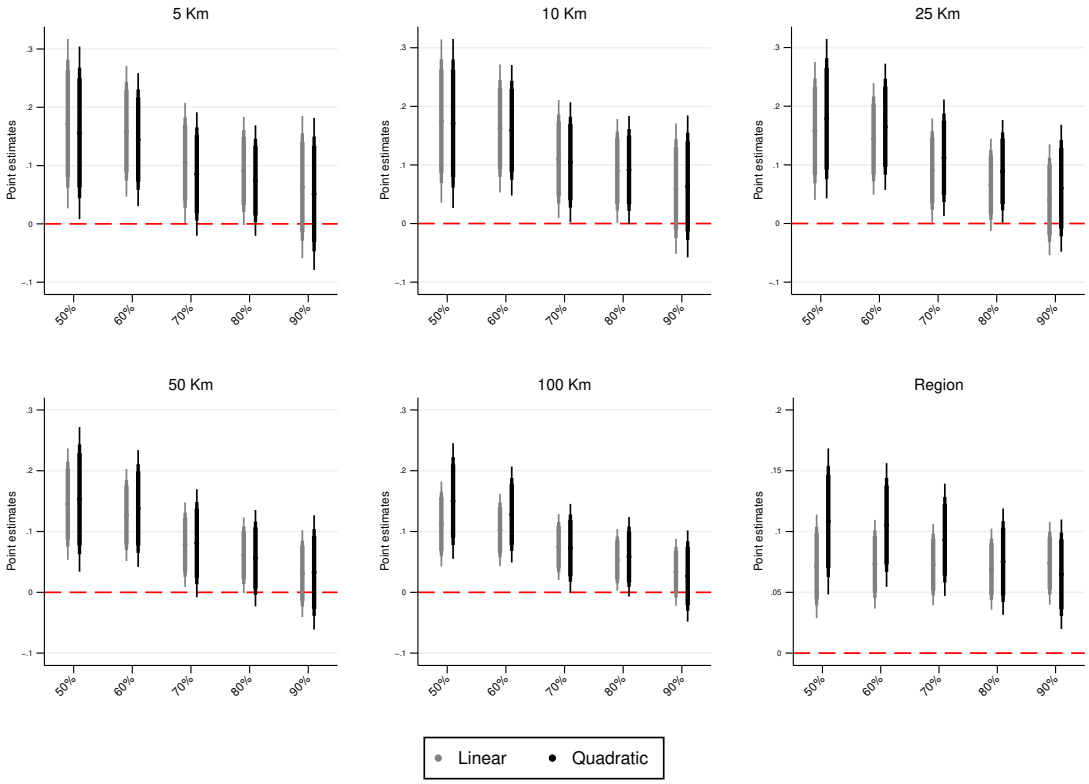
This plot shows the effect over time of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 4, with robust coefficients (Calonico et al., 2014, 2019, 2018, 2020). The dependent variable is the predicted risk of infiltration at the municipal level. *Treated* × *2006* indicates the year prior to the arrival of the EU funding and the red vertical dashed line separates graphically the pre- and post-treatment periods. Each subplot focuses on different geographical distances from the border, from 5km to the entire region. The plot reports each point estimate for the effect of EU funds on the risk of infiltration in a specific year along with two different polynomial specifications (i.e., linear and quadratic) along with 90%, 95%, and 99% confidence intervals.

FIGURE S22. Event study results: main models, with mean predicted risk over 100 sensitivity models



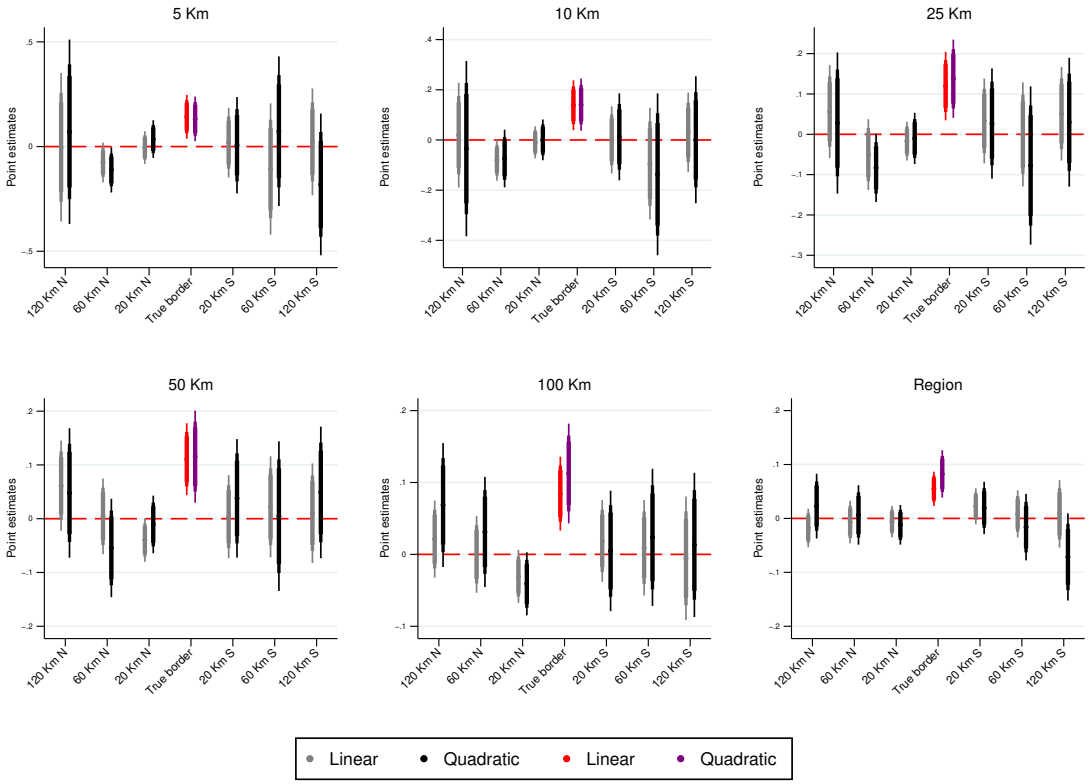
This plot shows the effect over time of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 4. The dependent variable is the average predicted risk of infiltration over 100 model runs with different combinations of random seeds. *Treated* × *2006* indicates the year prior to the arrival of the EU funding and the red vertical dashed line separates graphically the pre- and post-treatment periods. Each subplot focuses on different geographical distances from the border, from 5km to the entire region. The plot reports each point estimate for the effect of EU funds on the risk of infiltration in a specific year along with two different polynomial specifications (i.e., linear and quadratic) along with 90%, 95%, and 99% confidence intervals.

FIGURE S23. Difference-in-Discontinuities results for different thresholds and different discretization criteria



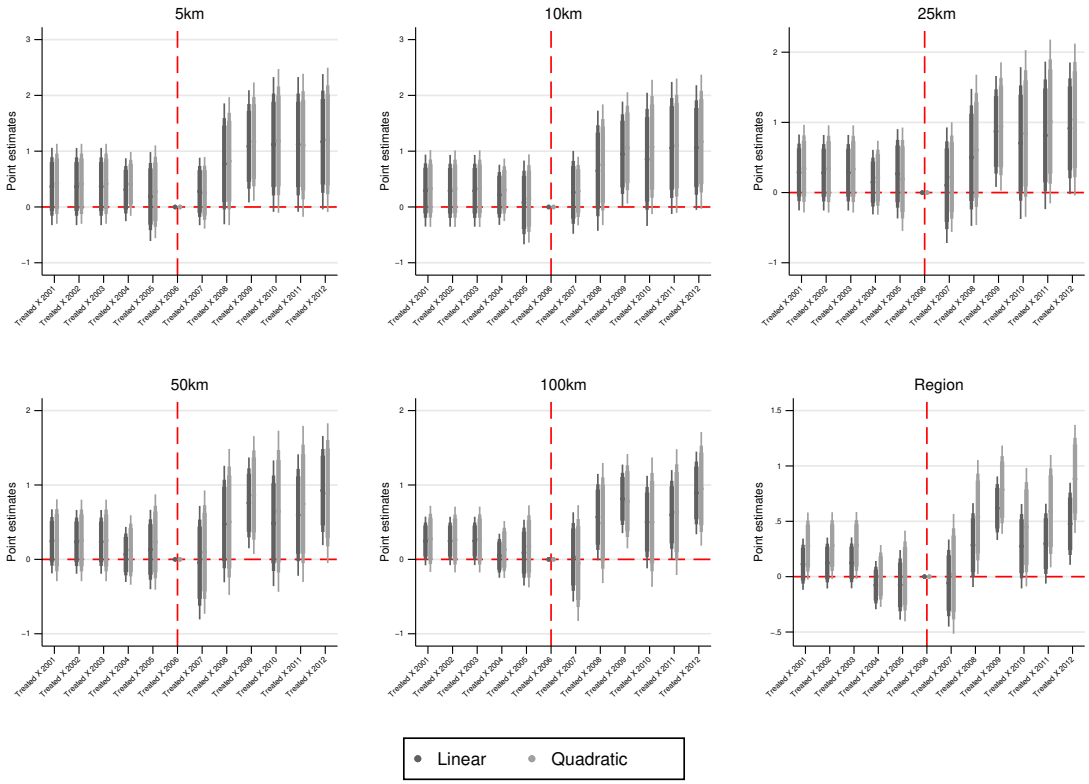
Note: This plot shows the effect of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 3. The dependent variable is a dummy variable that is equal to 1 if the predicted risk of infiltration at the municipal level is greater than a certain threshold, i.e., 50, 60, 70, 80, and 90%. Each plot represents a different bandwidth around the border. Gray lines refer to the specification using a linear polynomial for the distance to the border, while the black lines use a quadratic polynomial specification. The spikes for each line, from the thickest to the thinnest represent confidence intervals at 90, 95, and 99%, respectively.

FIGURE S24. Placebo tests: setting fake borders for different thresholds



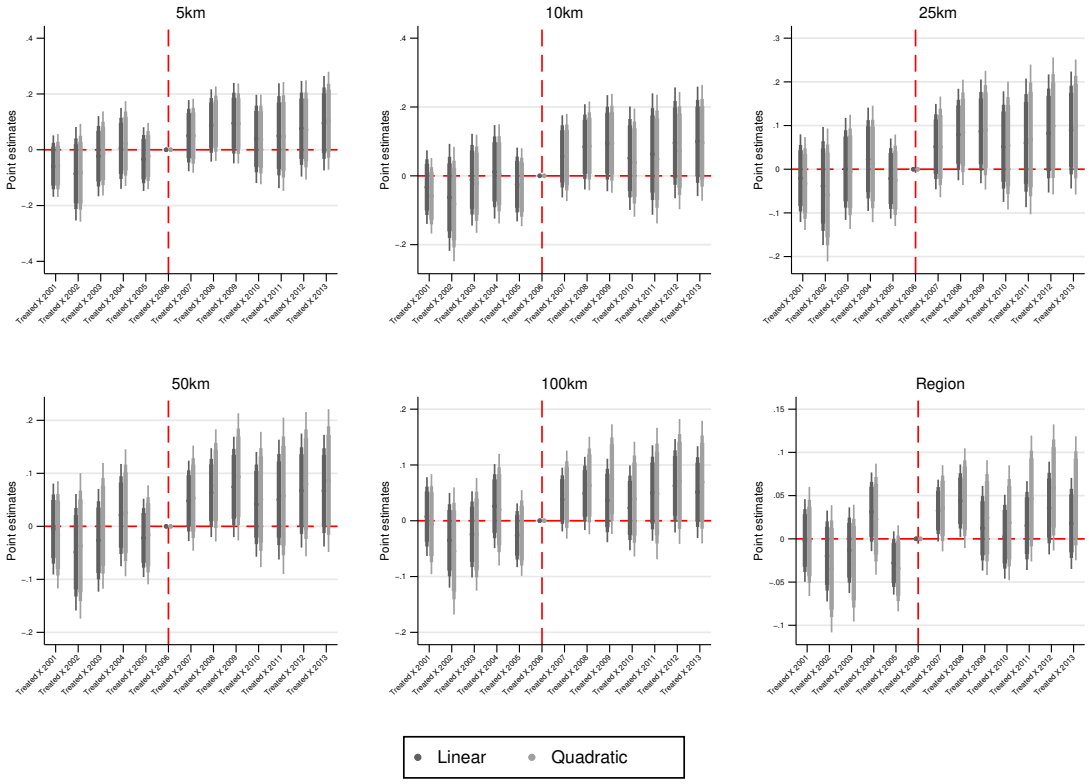
Note: This plot shows the effect of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 3. The dependent variable is the predicted risk of infiltration at the municipal level. The estimates are obtained by using either the true border between treated and control regions (i.e., red and purple lines) or fake borders constructed by moving the true one both north and south by 20, 60, and 120 km. Each plot represents a different bandwidth around the true or fake borders. Gray and red lines refer to the specification using a linear polynomial for the distance to the border, while the black and purple lines use a quadratic polynomial specification. The spikes for each line, from the thickest to the thinnest represent confidence intervals at 90, 95, and 99%, respectively.

FIGURE S25. EU Funds and Mafia-related Crimes – Event study



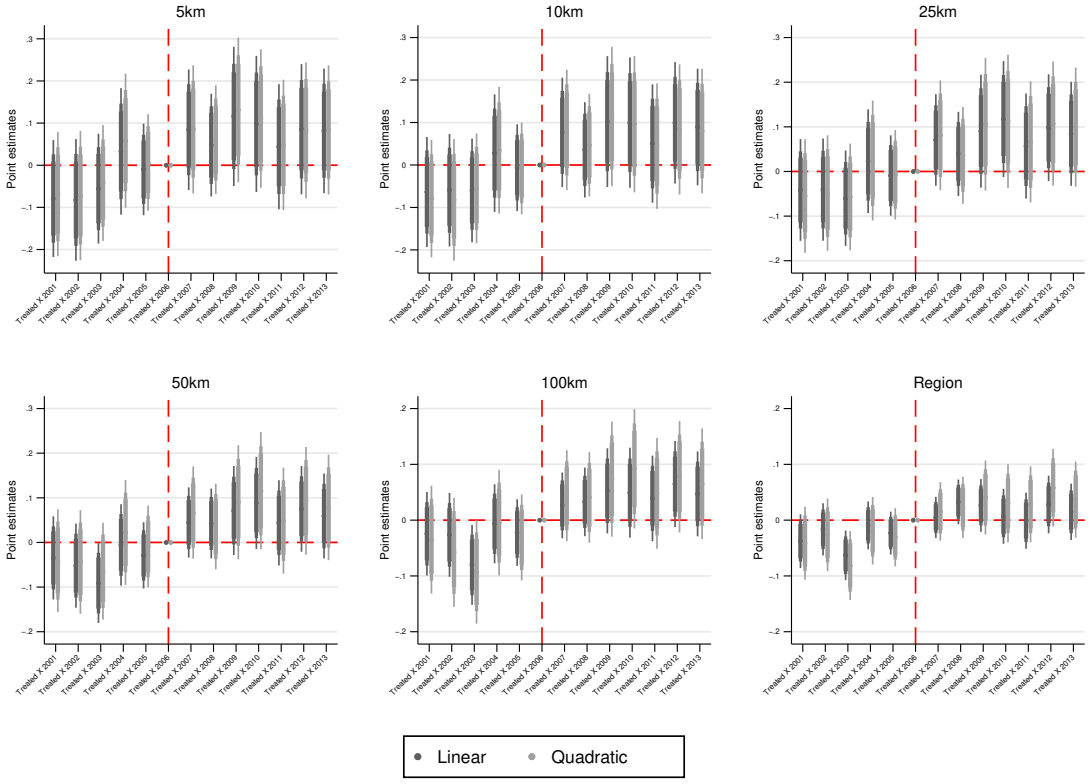
Note: This plot shows the effect over time of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 4 on the number of mafia-related crimes. The dependent variable is the count of mafia-related offenses at the municipal level. Data only refer to the 2004-2013 period, as no data on crimes were available before and after that period. *Treated* × *2006* indicates the year before the arrival of EU funding, and the red vertical dashed line graphically separates the pre- and post-treatment periods. Each subplot focuses on different geographical distances from the border, from 5km to the entire region. The plot reports each point estimate for the effect of EU funds on the number of mafia-related crimes in a specific year along with two different polynomial specifications (i.e., linear and quadratic) along with 90%, 95%, and 99% confidence intervals.

FIGURE S26. EU Funds and Risk of Mafia Infiltration – Event study (Index without Capital Expenditures)



Note: This plot shows the effect over time of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 4. The dependent variable is the predicted risk of infiltration at the municipal level trained without capital expenditure features. Treated × 2006 indicates the year prior to the arrival of the EU funding and the red vertical dashed line separates graphically the pre- and post-treatment periods. Each subplot focuses on different geographical distances from the border, from 5km to the entire region. The plot reports each point estimate for the effect of EU funds on the risk of infiltration in a specific year along with two different polynomial specifications (i.e., linear and quadratic) along with 90%, 95%, and 99% confidence intervals.

FIGURE S27. EU Funds and Risk of Mafia Infiltration – Event study (Index without Contemporaneous Spending Variables)



Note: This plot shows the effect over time of receiving the EU funds in the 2007-2013 wave obtained by estimating Equation 4. The dependent variable is the predicted risk of infiltration at the municipal level trained without contemporaneous public spending features. Treated × 2006 indicates the year prior to the arrival of the EU funding and the red vertical dashed line separates graphically the pre- and post-treatment periods. Each subplot focuses on different geographical distances from the border, from 5km to the entire region. The plot reports each point estimate for the effect of EU funds on the risk of infiltration in a specific year along with two different polynomial specifications (i.e., linear and quadratic) along with 90%, 95%, and 99% confidence intervals.

B.2 Additional Tables

TABLE S8. Difference-in-Discontinuities Results: main models, with robust coefficients

	(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	0.132***	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

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3, with robust coefficients (Calonico et al., 2014, 2019, 2018, 2020). The dependent variable is the predicted risk of infiltration. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the post-treatment period. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S9. Difference-in-Discontinuities Results: Mean Predicted Probabilities over 100 Sensitivity Models

	(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

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3. The dependent variable is the average predicted risk of infiltration over 100 model runs with different combinations of random seeds. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the post-treatment period. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects.Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S10. Difference-in-Discontinuities Results: predicted risk computed without Y=1 in the 2007-2013 frame

	(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

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3. The dependent variable is the predicted risk of infiltration computed after training a machine learning model that excluded observations with Y=1 in the 2007-2013 period. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the post-treatment period. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S11. Balancing Test

	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

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on different time-varying observable characteristics obtained by estimating Equation 3. The dependent variable are the value of several municipal features recorded both at 2001 and 2011. The coefficients shown are those for the interaction *Treat X Post 2006*, that is between being in a treated region (i.e., Puglia or Campania) with an indicator variable mapping the post-treatment period. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S12. Public spending features distribution change pre- and post-treatment

Regions	Type	N Features w/Sig. Change	% Tot. PS Features
Puglia and Campania	Treatment	36	40.00
Puglia	Treatment	34	37.78
Campania	Treatment	36	40.00
Sardinia	Placebo	0	0.00

Note: This table shows how many public spending features have changed their distribution in the post-treatment period. Significance is assessed via Mann-Whitney U Test. In total, 90 features are analyzed. Pre-treatment period considers the 2000-2005 frame. Post-treatment period considers the 2006-2013 frame.

TABLE S13. Difference-in-Differences results using Sardinia as placebo (given presence of EU funding but no presence of mafia)

	(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

Note: This table shows the analysis on the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating a difference-in-differences equation where the treated region is Sardinia and the control regions are Lazio and Molise. The dependent variable is the predicted risk of infiltration obtained via the best-performing machine learning model. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Sardinia) with an indicator variable mapping the post-treatment period. The models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S14. Difference-in-Discontinuities results with Risk Index trained without using Capital Expenditures

	(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	0.103**	0.096**	0.098**	0.102**	0.081**	0.096**	0.072***	0.078**	0.053***	0.073***	0.028**	0.053***
	(0.041)	(0.041)	(0.039)	(0.041)	(0.033)	(0.038)	(0.025)	(0.033)	(0.019)	(0.026)	(0.012)	(0.017)
Observations	1,298	1,298	2,023	2,023	3,931	3,931	7,094	7,094	10,885	10,885	17,086	17,086
R-squared	0.707	0.710	0.717	0.718	0.751	0.751	0.806	0.806	0.840	0.840	0.819	0.819
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	.0506	.0589	.0446	.0449	.0347	.0366	.0352	.0353	.0366	.0375	.0453	.0463
Bootstrap	.0119	.0211	.0125	.0136	.0131	.011	.0053	.0212	.0071	.0066	.0159	.0009
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	.1	.1

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3. The dependent variable is the predicted risk of infiltration obtained via the best-performing machine learning model, trained without features using the composition of capital expenditures. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the post-treatment period. The models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S15. Difference-in-Discontinuities results with Risk Index trained without using Contemporaneous Spending Variables

VARIABLES	(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	0.112***	0.106**	0.106***	0.110***	0.101***	0.111***	0.096***	0.098***	0.071***	0.100***	0.041***	0.066***
	(0.040)	(0.041)	(0.038)	(0.040)	(0.032)	(0.037)	(0.025)	(0.032)	(0.019)	(0.026)	(0.012)	(0.017)
Observations	1,298	1,298	2,023	2,023	3,923	3,923	7,033	7,033	10,774	10,774	16,968	16,968
R-squared	0.723	0.725	0.729	0.729	0.754	0.754	0.806	0.806	0.840	0.840	0.820	0.820
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	.0681	.0748	.0592	.0595	.0468	.0479	.045	.045	.0451	.0467	.0527	.0539
Bootstrap	.0058	.0105	.006	.0065	.0019	.0035	0	.0025	.0006	.0003	.0009	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	.09	.09	.1	.1	.09	.09

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3. The dependent variable is the predicted risk of infiltration obtained via the best-performing machine learning model, trained without contemporaneous public spending features. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the post-treatment period. The models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S16. Difference-in-Discontinuities results using count of mafia-related crimes as dependent variable (one unique post-treatment period)

	(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 Post 2006	0.474**	0.434*	0.456**	0.489**	0.305	0.383*	0.274*	0.310	0.311***	0.249	0.229***	0.298***
	(0.223)	(0.236)	(0.221)	(0.224)	(0.199)	(0.220)	(0.159)	(0.200)	(0.119)	(0.165)	(0.073)	(0.103)
Observations	1,198	1,198	1,867	1,867	3,621	3,621	6,509	6,509	9,964	9,964	15,679	15,679
R-squared	0.388	0.389	0.374	0.375	0.359	0.359	0.364	0.366	0.394	0.394	0.405	0.405
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	.1858	.187	.1777	.1794	.1687	.1693	.176	.1781	.2036	.2037	.2099	.2103
Bootstrap	.0377	.0719	.0378	.0281	.1278	.0849	.0889	.1312	.0071	.1329	.0015	.0032
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	.86	.86	.86	.86	.8	.8	.84	.84	.74	.74	.66	.66

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on mafia crimes obtained by estimating Equation 3. The dependent variable is the count of mafia-related crimes. Observations refer to the 2004-2013 period. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the post-treatment period. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S17. Difference-in-Discontinuities results using count of mafia-related crimes as dependent variable (two separate post-treatment period)

	(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	0.260	0.215	0.261	0.281	0.092	0.178	0.064	0.112	0.155	0.034	0.079	0.142
	(0.234)	(0.247)	(0.272)	(0.238)	(0.286)	(0.285)	(0.248)	(0.289)	(0.187)	(0.254)	(0.119)	(0.166)
Treat X 2009 - 2012	0.858***	0.844**	0.786**	0.860***	0.615**	0.723**	0.540***	0.598**	0.556***	0.531**	0.383***	0.536***
	(0.317)	(0.338)	(0.304)	(0.317)	(0.263)	(0.299)	(0.201)	(0.263)	(0.149)	(0.209)	(0.086)	(0.123)
Observations	1,198	1,198	1,867	1,867	3,621	3,621	6,509	6,509	9,964	9,964	15,679	15,679
R-squared	0.399	0.402	0.382	0.385	0.364	0.365	0.369	0.370	0.397	0.398	0.406	0.407
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	.2001	.2035	.1879	.192	.1757	.1772	.182	.184	.2076	.2084	.2117	.2127
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	.86	.86	.86	.86	.8	.8	.84	.84	.74	.74	.66	.66

Note: This table shows the analysis of the effect of receiving EU funds in the 2007-2013 wave on mafia crimes obtained by estimating Equation 3. The dependent variable is the count of mafia-related crimes. Observations refer to the 2004-2013 period. *Treat X 2007-2008* is the interaction between being in a treated region (i.e., Campania or Puglia) with an indicator variable mapping the first post-treatment period. *Treat X 2009-2012* is the interaction between being in a treated region and an indicator variable mapping the second post-treatment period. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S18. Difference-in-Discontinuities results to investigate mechanisms of increased mafia infiltration. Mechanism 1: Relaxation of EU monitoring rules

	(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	0.271**	0.246**	0.268**	0.270**	0.193*	0.246**	0.130	0.182*	0.095	0.118	0.108**	0.075
	(0.115)	(0.092)	(0.119)	(0.112)	(0.102)	(0.113)	(0.081)	(0.101)	(0.061)	(0.085)	(0.051)	(0.079)
Observations	247	247	387	387	790	790	1,683	1,683	3,111	3,111	5,629	5,629
R-squared	0.833	0.840	0.840	0.841	0.869	0.872	0.896	0.896	0.900	0.900	0.866	0.866
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	.2069	.239	.1794	.1843	.1215	.1371	.0914	.0995	.0571	.0619	.0462	.0471
Bootstrap	.0369	.039	.0448	.0388	.113	.0702	.1575	.1365	.1444	.2521	.0332	.4832
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

Note: This table shows the analysis on the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3. The dependent variable is the predicted risk of infiltration obtained via the best-performing machine learning model. *Treat X Post 2006* is the interaction between being in a treated region (i.e., the part of Campania closer to Lazion than Molise) with an indicator variable mapping the post-treatment period. The analysis excludes Puglia and the part of Campania closer to Molise than Lazio as treated regions and Molise as a control region. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

TABLE S19. Difference-in-Discontinuities results to investigate mechanisms of increased mafia infiltration. Mechanism 2: Increased EU funding + Relaxation of EU monitoring rules

	(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	0.133***	0.127***	0.132***	0.132***	0.121***	0.134***	0.115***	0.119***	0.105***	0.123***	0.098***	0.123***
	(0.043)	(0.044)	(0.040)	(0.042)	(0.034)	(0.039)	(0.028)	(0.035)	(0.024)	(0.031)	(0.019)	(0.025)
Observations	1,051	1,051	1,636	1,636	3,133	3,133	5,350	5,350	7,663	7,663	11,339	11,339
R-squared	0.649	0.652	0.647	0.647	0.660	0.661	0.732	0.732	0.789	0.789	0.772	0.772
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	.0668	.0766	.0584	.0587	.046	.0482	.0443	.0448	.0486	.0491	.0574	.0581
Bootstrap	.0024	.0047	.0011	.0029	.0006	.0008	0	.0011	0	0	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	.21	.21	.18	.18	.13	.13	.12	.12	.12	.12	.12	.12

Note: This table shows the analysis on the effect of receiving EU funds in the 2007-2013 wave on the risk of mafia infiltration in local politics obtained by estimating Equation 3. The dependent variable is the predicted risk of infiltration obtained via the best-performing machine learning model. *Treat X Post 2006* is the interaction between being in a treated region (i.e., Puglia or the part of Campania closer to Molise than Lazio) with an indicator variable mapping the post-treatment period. The analysis excludes The part of Campania closer to Lazio than Molise as a treated region and Lazio as a control region. Columns (1) and (2) consider municipalities that are within 5km of the border separating treated and control regions. Columns (3) and (4) expand the sample to municipalities up to 10 km. Following, columns (5) and (6) focus on the 25km distance, (6) and (7) on the 50km distance, (8) and (9) on 100km and (10) and (11) consider the entire region. We fit two polynomial specifications per distance threshold: linear and quadratic. All models include municipality and year-fixed effects. Standard errors are adjusted for serial correlation, and the table also presents the p-value obtained with standard errors bootstrapped 10,000 times. The last row presents the baseline value of the dependent variable for control municipalities as of 2006. *, **, *** indicate statistical significance at the conventional 10%, 5% and 1% levels, respectively.

References - Supplementary Materials

Batista, Gustavo E. A. P. A., Ronaldo C. Prati, and Maria Carolina Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD Explorations Newsletter*, June 2004, *6* (1), 20–29.

Bishop, Christopher M., *Neural Networks for Pattern Recognition*, Clarendon Press, November 1995.

Breiman, Leo, "Random Forests," *Machine Learning*, October 2001, *45* (1), 5–32.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell, "On the effect of bias estimation on coverage accuracy in nonparametric inference," *Journal of the American Statistical Association*, 2018, *113* (522), 767–779.

_ , _ , and _ , "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs," *The Econometrics Journal*, 2020, *23* (2), 192–210.

_ , Matias D. Cattaneo, and Rocio Titiunik, "ROBUST NONPARAMETRIC CONFIDENCE INTERVALS FOR REGRESSION-DISCONTINUITY DESIGNS," *Econometrica*, 2014, *82* (6), 2295–2326.

_ , Matias D Cattaneo, Max H Farrell, and Rocio Titiunik, "Regression discontinuity designs using covariates," *Review of Economics and Statistics*, 2019, *101* (3), 442–451.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, June 2002, *16*, 321–357.

Chen, Tianqi and Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System," in "Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining" KDD '16 Association for Computing Machinery New York, NY, USA August 2016, pp. 785–794.

Colonnelli, Emanuele, Jorge A. Gallego, and Mounu Prem, "What Predicts Corruption?," *SSRN Electronic Journal*, 2020. Available at SSRN 3330651.

Friedman, Jerome H., "Greedy Function Approximation: A Gradient Boosting Machine," *The Annals of Statistics*, 2001, *29* (5), 1189–1232. Publisher: Institute of Mathematical Statistics.

Haibo He, Yang Bai, Edwardo A. Garcia, and Shutao Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in "2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)" IEEE Hong Kong, China June 2008, pp. 1322–1328.

Hoerl, Arthur E. and Robert W. Kennard, "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, 2000, *42* (1), 80–86. Publisher: [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality].

Lundberg, Scott M. and Su-In Lee, "A Unified Approach to Interpreting Model Predictions," November 2017. arXiv:1705.07874 [cs, stat].

_ , Gabriel G. Erion, and Su-In Lee, "Consistent Individualized Feature Attribution for Tree Ensembles," March 2019. arXiv:1802.03888 [cs, stat].

Quinlan, J. R., "Induction of decision trees," *Machine Learning*, March 1986, *1* (1), 81–106.

Shapley, Lloyd S., "A Value for N-Person Games," Technical Report, RAND Corporation March 1952.

Tibshirani, Robert, "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 2011, *73* (3), 273–282. Publisher: [Royal Statistical Society, Wiley].

Zou, Hui and Trevor Hastie, "Regularization and Variable Selection Via the Elastic Net," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, April 2005, *67* (2), 301–320.