

# When Nation Building Goes South: Draft Evasion, Government Repression, and the Origins of the Sicilian Mafia

Gianni Marciante\*  
University of Warwick

This version: February 27, 2023

## Abstract

Can coercive nation building foster organised crime? I provide evidence on this question by showing how a coercive nation-building policy – conscription – contributed to the emergence of the Sicilian mafia in the 19th century. This policy generated widespread demand for private protection in Sicilian society, especially after a government-led repression campaign run in 1863 to curb large-scale draft evasion on the island. Using a novel dataset on episodes of government repression and historical mafia presence, I find that the Sicilian mafia was more likely to develop in towns victimised by the repression campaign. To infer causality, I use an instrumental variables approach based on a network of least-cost paths connecting garrison towns that were visited by the army during the 1863 expedition. I find suggestive evidence that increasing distrust in government in repressed towns helps account for the early spread of the Sicilian mafia.

**Keywords:** Nation Building, Conscription, Repression, Organised Crime, Mafia.

**JEL Codes:** K42, H11, P16.

---

\*I am very grateful to James Fenske and Sharun Mukand for their guidance and support during this project. For very helpful comments and discussions, I also thank Brian A’Hearn, Luca Bagnato, Sascha Becker, Alberto Bisin, Tancredi Buscemi, Felipe Campante, Stefano Caria, Carlo Ciccarelli, Giuseppe De Feo, Arcangelo Dimico, Stefano Fiorin, Ludovica Gazzè, Bishnupriya Gupta, Muhammad Haseeb, Michael Hauptert, Amrita Kulka, Eric Melander, Martina Miotto, Laura Ogliari, Claudia Rei, Ana Tur-Prats, Felipe Valencia Caicedo, Joachim Voth and Chenzi Xu, as well as audiences at Collegio Carlo Alberto, LSE, Manchester, Queen Mary, Southern Denmark, Tor Vergata, Warwick, Zurich, the 47th Spanish Economic Association Simposio, the 2022 ASREC Europe Conference, the 34th EAEPE Conference, the XIX World Economic History Congress, the 14th EHES Conference, the 2022 Cliometric Conference, the 2022 EHS Annual Conference and the 2021 EHS PhD Thesis Workshop. I am indebted to the archivists and librarians at the National Museum of the Italian Risorgimento, the Historical Archive of the Chamber of Deputies, and the State Archives of Agrigento, Caltanissetta, Enna, Messina, Palermo and Trapani for their critical assistance. Veronica Verdesca provided exceptional research assistance. Financial support from the Royal Economic Society is gratefully acknowledged. All mistakes are my own.

# 1 Introduction

Political leaders, with Napoleon as an early example, have often viewed nation-building policies as a key instrument to increase cultural homogeneity and improve political stability (Aghion et al., 2019; Alesina et al., 2021). Cultural homogenisation policies continue to attract attention in many countries today, as governments seek to bridge social cleavages that may otherwise generate social unrest and lead to collective action failures (Alesina et al., 1999; Miguel and Gugerty, 2005). The correct design of these policies is therefore of paramount importance for policy makers.

Recent work by Fouka (2020) and Dehdari and Gehring (2022), among others, argues that *coercive* nation-building policies lacking appropriate economic incentives may bring cultural resistance by minority communities.<sup>1</sup> Nevertheless, little is known about the specific local conditions that facilitate or hinder these acts of resistance.

In this paper, I partly fill this gap by focusing on one possible condition that could encourage cultural resistance: the presence of informal social institutions supplying private protection. Members of the minority targeted by the forced assimilation reforms may decide to evade these policies in order to protect their “oppositional” identity (Scott, 2009; Bisin et al., 2011; Carvalho, 2013; Carvalho and Koyama, 2016). In doing so, they might turn to local agents to seek protection from government law enforcers, in order to preserve their distinctive cultural heritage. This evasion could be more likely to occur in the presence of stronger homogenisation policies. This resistance process could eventually produce two results: lower social integration and the rise of mafia-type organisations supplying protection to people unwilling to integrate.

---

<sup>1</sup>Dehdari and Gehring (2022) distinguish between benevolent and coercive nation-building policies. The first includes public investments in public goods, such as roads or railways, through which a government can promote social connectedness. Coercive forms of nation building, on the other hand, restrict personal and/or political freedoms, and range from language bans and censorship to military conscription and segregation. Violent homogenisation policies have been recently seen in Russia’s invasion of Ukraine and in the Xinjiang region in China (Human Rights Watch, 2021; Said-Moorhouse and Ochman, 2022).

I test the above hypothesis in the context of 19th-century Sicily. Combining historical data from several archival sources, I show how one coercive nation-building policy – military conscription in the 1860s – provoked an identity backlash that contributed to the early spread of the Sicilian mafia. From its establishment in 1861, the Kingdom of Italy’s central government implemented a range of nation-building policies, including conscription, to strengthen internal stability (Duggan, 2007; Alesina et al., 2021). The introduction of conscription in Sicily met strong local resistance, leading many young men to avoid the draft and join the ranks of local bandits (Carraro, 1976; Riall, 1998). The central government’s response to the high rates of draft evasion was a campaign of military repression from June to November 1863, to round up Sicilian deserters and draft evaders. This campaign increased distrust in the government and, in turn, helped the Sicilian mafia gain popular support in communities directly affected by these repressive measures (Alatri, 1954; Dickie, 2004). Moreover, fear of the bandits still hiding in the Sicilian countryside and of the law enforcement officials searching for draft evaders generated widespread demand for private protection in repressed towns. This demand for protection was met by the mafia, the sole remaining trusted provider in the protection industry after 1863.

I construct a novel dataset on town-level exposure to the 1863 military campaign, assembled largely from archival material including handwritten military reports and private letters written by Giuseppe Govone, the army general in charge of the military operation. To identify variation in exposure to government repression again at the town level, I further digitise three Sicilian newspaper archives reporting local daily news from June to December 1863.<sup>2</sup> The information on these sources is then combined with newly digitised data on historical mafia presence, from the Borsani-Bonfadini parliamentary enquiry on the conditions of public security in

---

<sup>2</sup>I focus on the following historical daily newspapers: *Il Giornale Ufficiale di Sicilia*, *Il Precursore* and *La Forbice*. See Section 3.2 for more details.

Sicily, conducted between 1874 and 1876. I leverage the information in this enquiry to infer mafia presence at the town level in 1875.

The ordinary least squares (OLS) estimates indicate that the Sicilian mafia was more likely to develop in towns victimised by the 1863 repression campaign. This finding is consistent with anecdotal evidence suggesting a rise in demand for protection from government officials by the Sicilian population living in repressed towns (Alongi, 1977). Alternative explanations for my results are, however, possible. The army could have concentrated its efforts on places where the mafia had already taken root, and so differences in exposure to repression may simply pick up the effect of pre-existing criminal activity across towns. At the same time, other factors, such as trust in the government, may correlate with exposure to the military campaign (as places showing distrust in government could be more prone to draft evasion) and subsequent mafia diffusion. Selection into repression could also be negative, as the military may have targeted towns with a low crime rate to exert less effort in the search for draft evaders. I adopt a number of empirical strategies to address these concerns.

First, I control for town-level banditry presence before the Italian unification of 1861, to capture local differences in pre-repression illegal activity. Second, I account for ex-ante local attitudes towards the central government and also control for a large set of observables that proxy for economic activity and local state presence. Third, I include province fixed effects to control for unobserved confounders that are common across towns within small geographical areas. Estimates of Oster (2019)'s degree of selection on unobservables relative to observables also suggest that my result is unlikely to be largely attributed to unobserved heterogeneity at the local level.

Finally, I use an instrumental variables (IV) strategy based on the “inconsequential places” approach of Redding and Turner (2015) to estimate the

causal effect of exposure to government repression on early mafia presence.<sup>3</sup> I exploit archival evidence about the army’s itinerary to construct least-cost paths connecting garrison towns that were visited by the soldiers, using the pre-existing (1826) road network.<sup>4</sup> I then calculate the distance from these least-cost paths to each town in the dataset, and I use this distance as an instrument for exposure to repression.<sup>5</sup> Conditioning on the distance to the main commercial roads in my IV analysis provides additional support for the instrument’s exogeneity. The two-stage least squares (2SLS) estimates confirm the OLS results, suggesting that violent exposure to a coercive nation-building policy was a key driver of the Sicilian mafia’s early appearance. The 2SLS results indicate that exposure to the repression campaign increased the probability of mafia presence by approximately 79 percentage points, compared to a mean of 10.6% in the restricted sample without garrison towns.

My findings persist if different transformations of the excluded instrument, such as inverse hyperbolic sine or binarisation, are employed. Similarly, replicating the analysis using different estimation techniques such as propensity score matching or geographical matching delivers comparable results. I additionally perform three falsification exercises to rule out any direct effect of my excluded instrument on mafia presence in 1875. In these falsification tests I do not find any significant correlation between the distance from the network of least-cost paths and proxies for pre-repression criminal activity and trust in institutions.

I explore two mechanisms that could drive the results: increasing demand for mafia by the population in repressed towns, and rising mafia supply in the same areas. Towns exposed to government repression may have experienced a decline in

---

<sup>3</sup>Other recent uses of this approach include Hornung (2015), Fenske et al. (2021) and Bogart et al. (2022).

<sup>4</sup>It is highly likely that existing garrison towns in 1863 Sicily played an important role in determining the itinerary of Govone’s troops, since in that period military companies used to leave on expedition from these locations and finally return to them once their military operations were concluded (Stato Maggiore del Regio Esercito, 1863).

<sup>5</sup>I finally remove the visited garrison towns from the sample, to identify the effect of repression in towns that were only treated incidentally due to their position along the least-cost path between two garrison towns. See Section 4.3 for more details on this identification strategy.

state legitimacy after 1863, which could have induced more people to rely on the local *mafiosi* to get protection. At the same time, the repression campaign could have triggered an increase in the number of draft evaders in repressed towns. By lowering the opportunity cost of evading the draft through the supply of private protection, the Sicilian mafia could have benefited from the increase in the criminal labour force and hired from a larger pool of outlaws.

To investigate the demand-side mechanism, I first collect new data on the general elections held from 1861 to 1867 from the historical archive of the Chamber of Deputies of the Italian Parliament. I compute voter turnout at the town level and use it as a proxy for trust in institutions (Guiso et al., 2000; Cannella et al., 2021). The relationship between government repression in 1863 and pre-existing trust in institutions among the Sicilian electorate is not significant. In contrast, I find that eligible voters in repressed towns were less likely to vote in 1867, which is consistent with the increasing-distrust-in-government argument proposed above. Using another proxy for trust in the central government, smallpox vaccination uptake, I also show that towns exposed to government repression exhibited systematically lower vaccination rates in the period 1880-1884 relative to non-repressed towns.<sup>6</sup>

Focusing on the supply-side mechanism, I use an event study on newly assembled data on department-level draft evasion rates before and after 1863. I find little empirical evidence that the 1863 repression campaign led to a higher supply of criminals in the affected areas. Taken together, the evidence suggests that state repression generated distrust in the central government, and this helped the Sicilian mafia win broad support and extend its influence. The repression-induced political dissent, however, did not significantly affect the number of people participating in local criminal activity.

I conclude by showing that exposure to the 1863 military repression is also a

---

<sup>6</sup>In Section 5.1 I explain why lower vaccine take-up rates in this context may be more indicative of local distrust in government than of skepticism about the smallpox vaccine.

positive and significant predictor of mafia presence in 1900, using town-level data on mafia prevalence from Cutrera (1900). Overall, my results suggest that episodes of government repression during historical critical junctures can generate enduring social consequences.

## 1.1 Contribution

This paper speaks to several existing literatures. First, it extends the literature on the causes and consequences of culture (e.g. Fernández, 2011; Gorodnichenko and Roland, 2011; Desmet et al., 2017), focusing on the interplay between culture and institutions (Alesina and Giuliano, 2015; Lowes et al., 2017; Bisin and Verdier, 2017). Particularly relevant to my context is the contribution of Acemoglu and Robinson (2021), who emphasize the fluidity of cultures, to be intended as a set of cultural attributes that can be rewired as conditions change. The anecdotal evidence presented in this paper suggests that 19th-century Sicily was characterised by a relatively “hardwired” culture that limited the formation of a new cultural configuration in response to the Italian unification process.

This paper also relates to the literature on nation building, i.e. policies and events that affect the construction of national identity. Several scholars have analysed the effects of homogenisation policies and historical episodes influencing the sense of national identity, finding mixed results. On the one hand, a remarkable body of work has documented a larger adoption of the national identity at the local level, following the introduction of language unification reforms (You, 2018), population resettlement programs in fractionalised communities (Bazzi et al., 2019), exposure to government propaganda (Blouin and Mukand, 2019) and shared experiences (Depetris-Chauvin et al., 2020; Esposito et al., 2021), and the activity of individual leaders (Assouad, 2020). On the other hand, nation-building attempts by national governments have also been shown to lead to lower cultural homogenisation, as seen, for instance, with German immigrants in the US after World War I (Fouka, 2020) or in Catalonia

after the introduction of bilingualism in 1983 (Clots-Figueras and Masella, 2013).<sup>7</sup> Scholars have also studied nation-building policies as a possible outcome variable, in response to mass migration (Bandiera et al., 2019), the threat of external wars (Aghion et al., 2019; Alesina et al., 2020), and the risk of overthrow of non-democratic rulers (Alesina et al., 2021).

I add to this literature in two ways. First, I focus on the impact of military conscription, an under-studied but important nation-building policy, and investigate the consequences of resistance to it by the targeted social group.<sup>8</sup> Second, and more importantly, I address a specific form of nation building implemented through coercion, and analyse its repercussions for the protection market.<sup>9</sup> With the focus on coercive nation building and state repression, my work builds on research by Dehdari and Gehring (2022). These authors exploit the quasi-exogenous division of the historically homogeneous French regions of Alsace and Lorraine in 1871 to provide evidence of the effect of exposure to forced nation building on identity formation. They document a strengthening of regional rather than national identity expressed by the treated community, but their work does not discuss any impact of forced assimilation policies on the level of criminal activity.

This paper also contributes to the growing literature on the origins of organised crime, focusing on the historical evolution of the Sicilian mafia. Seminal works by leading mafia experts such as Gambetta (1993) and Lupo (1996) have highlighted the role of mafia-type organisations as suppliers of private protection, particularly in societies with low trust and weak democracy. More recently, a number of studies have stressed specific features of 19th-century Sicily that may have helped the Sicilian

---

<sup>7</sup>In the same vein, Carvalho and Koyama (2016) show through a theoretical model that, when facing education policies aiming at transmitting the majority culture, historically marginalised communities can underinvest in education as a form of cultural resistance.

<sup>8</sup>More recent examples of conscription programmes structured to promote a sense of national identity occurred in Ghana, Iraq, Nigeria and South Korea, with discordant results. See Shurkin et al. (2017) for an overview of these case studies.

<sup>9</sup>On the political consequences of state repression, see, for instance, Rozenas et al. (2017), Rozenas and Zhukov (2019), Komisarchik et al. (2020) and Cannella et al. (2021).



mafia to spread, such as the high level of land fragmentation (Bandiera, 2003), the profitability of citrus production (Dimico et al., 2017) and sulphur extraction (Buonanno et al., 2015; Ciccarelli et al., 2021), or the harsh working conditions of the island's peasants (Acemoglu et al., 2020). All of these factors generated demand for private protection that was met by the mafia, in the absence of strong formal institutions in Sicily. I add to this body of work by looking at the direct role of the Italian government in fostering the early development of the Sicilian mafia at the time of Italy's unification. Moreover, I empirically document the initial rise of the Sicilian mafia as a result of lack of state legitimacy, due to the violent exposure to a coercive nation-building policy.

Lastly, this study draws on the literature on the consequences of radical and externally imposed reforms, in the context of post-unification Italy.<sup>10</sup> The policy reforms implemented in Southern Italy by the Piedmontese elites (who were from Northern Italy and led the unification process) soon after the establishment of the Kingdom of Italy in 1861, as well as the local resistance to them, have been researched extensively (Amodio, 2012; Lecce et al., 2022). More closely related to this paper is Accetturo et al. (2017), who exploit the adoption of the deeply repressive Pica Law in 1863 in the Italian South (excluding Sicily and Sardinia) to estimate the causal impact of tougher law enforcement on political participation. This paper makes several contributions to this literature. Not only do I focus on the role of organised crime, but I also investigate the effect of a government policy imported to Sicily, a region not yet studied in the rest of this literature. In addition, I provide novel evidence on the consequences of a different form of protest against the new institutional framework, namely draft evasion, following the introduction of a coercive nation-building policy.

---

<sup>10</sup>See Dryzek (1996), De Jong and Mamadouh (2002) and Roland (2004) for other studies on the consequences of institutional changes. See also Acemoglu et al. (2011), Kopsidis and Bromley (2016) and Lecce and Ogliari (2019) on the effect of the Napoleonic institutions on the invaded territories during the Napoleonic military campaigns.

This paper is structured as follows. In Section 2 I give a brief overview of the early development of the Sicilian mafia and discuss the events characterising the forced introduction of conscription in Sicily. Section 3 describes the historical data. Section 4 illustrates the empirical strategy and presents the main empirical findings. In Section 5 I provide evidence on mechanisms, while in Section 6 I present a persistence analysis. Section 7 concludes.

## 2 Historical Background

This section outlines the historical context in which the Sicilian mafia emerged. The historical analysis starts with a description of the status of public order in Sicily before the Italian unification. The discussion then moves towards the events related to the enactment of the conscription law in 1861 and the following decline in state legitimacy that provided fertile ground for the spread of the Sicilian mafia.

### 2.1 The Origins of the Sicilian Mafia

Although there are documented examples of mafia-like practices in Sicily as early as the end of the 16th century (Cancila, 1984), most works (e.g. Franchetti, 1877; Gambetta, 1993; Lupo, 1996; Dickie, 2004) identify the interval between the end of feudalism and the first decades after the Italian unification as the period of the mafia's first appearance. Before the abolition of feudal land relations in 1812, about three-quarters of Sicilian villages were under the feudal barons' authority (Bandiera, 2003). The distant governance of the Bourbon monarchy, ruling the island from 1735, resulted in few constraints on barons' activities within their fiefs. Importantly, most of these landlords, especially in Western Sicily, appointed private guards (named *bravi*, who were largely former bandits) to protect their land and keep peasants under control (Blok, 1975).

Three major events followed the demise of feudalism and the transformation of land into a market commodity in 1812. First, power conflicts between the barons and

the Spanish monarch intensified, leading to low levels of law enforcement and publicly provided security throughout Sicily, which lasted until the fall of the Bourbon Kingdom in 1860 (Mack Smith, 1968; Buonanno et al., 2015). Second, landowners' control over their lands slowly weakened, to the advantage of tenants (*gabelloti*), who saw an unprecedented opportunity for social improvement through the purchase of lands previously belonging to the baronial estates. Many *gabelloti* went from renters to proprietors and estate managers, while keeping the same production structure used by feudal lords. Third, once released from baronial control, the members of the barons' militias became autonomous suppliers of protection, offering their services to classes other than the aristocracy (Gambetta, 1993). Local groups of protectors (called *cosche*, whose members were known as *campieri*) emerged and started supporting "even small interests which formerly could count only on the muscle and energy of their bearers" (Franchetti, 1877, p. 159).

The local *cosche* used to maintain good relations with the Bourbon government's rural law enforcement agents (*compagnie d'armi*), since the latter were mainly former outlaws. A system of collusion between the *campieri*, the *compagnie d'armi* and the local brigands<sup>11</sup> developed, with the *campieri* manipulating the demand for protection by exploiting violent threats made by the local brigands at the *campieri*'s request, with the tacit consent of the law enforcers.<sup>12</sup> It was in this context that the first mafia-style aggregates, named *proto-mafia* by Lupo (1996), appeared.<sup>13</sup>

---

<sup>11</sup>As a consequence of the abolition of common rights, peasants and shepherds experienced extreme poverty after 1812 and many of them were pushed into banditry to survive (Gambetta, 1993). The largest increases in the incidence of banditry before the Italian unification were registered in 1820, 1837 and 1848 (Bandiera, 2003).

<sup>12</sup>The first to detect this perverse system at the time was Attorney General Pietro Calà Ulloa in 1838, when he denounced the presence of "brotherhoods" in the towns, acting as "little governments within the government" (Ulloa, 1838). The link between banditry and mafia is also highlighted by Cutrera (1900, pp. 90-91) as follows: "Banditry per se has nothing to do with the mafia, but it contributes to make mafia powerful, putting itself at mafia's service. [...] If the big shots in the province provide brigands with help and protection, it is [...] simply because, being mafia bosses, they increase their moral influence, and so their omnipotence, by keeping brigands at their service."

<sup>13</sup>See also Colajanni (1900), Chapter 4, for a discussion of early mafia presence in Sicily under Bourbon rule.

This system continued even after Italy’s unification in 1861 and found new strength in the discontent towards the newly formed Italian government, which has endured ever since.

## 2.2 Italian Unification and Coercive Nation Building in Sicily

The years immediately after the Italian independence war of 1859 were turbulent for the new Kingdom of Italy. In 1860 General Giuseppe Garibaldi undertook the *Expedition of the Thousand* to conquer the Kingdom of the Two Sicilies and achieve a united Italy. This would set the stage for the imposition of Northern Italy’s institutions (and those of the Kingdom of Sardinia-Piedmont to a large extent) onto the annexed territories of the South (Clark, 2009; Lecce et al., 2022). Starting from 1861, the entire Piedmontese institutional framework, including the whole administrative and judicial system, was extended to Southern Italy. Most of the reforms that the central government pursued in the South soon after unification were driven by the immediate need for institutional unity, leaving little room to accommodate the specific needs of the newly conquered regions (Riall, 1998; Pezzino and Mangiameli, 2016). Any attempt to improve the archaic road network, the poor educational system or the inefficient land distribution characterising Southern Italy was made subservient to the overriding goal of giving the new state solid institutional foundations (Riall, 1998, p. 154). The new central government’s political approach triggered a wave of popular unrest especially in the Southern mainland, where it became known as *brigandage* (Lecce et al., 2022).

Facing mounting internal instability, the central government decided on a set of nation-building policies to reduce political tensions and “make the Italians.” Compulsory schooling and mandatory military service were the main forces used for unifying the population. For the new rulers linguistic homogenisation became essential for the construction of a national identity (Alesina et al., 2021). In the same

vein, the national army soon developed into “the true school of the nation” (Ministero della Guerra, 1872) and “the steel thread that had sewn Italy and was keeping it united,” using the words of the veteran Risorgimento patriot Luigi Settembrini.<sup>14</sup>

These policies quickly came to be perceived as unpopular obligations imposed by the ruling elite. The introduction of conscription in Sicily, in particular, was met with strong local resistance. Before the Italian unification, there was no draft on the island. When this policy was implemented in 1861, each year 20-year-old men were called up to serve up to five years in the military, far from home and alongside culturally different people from other Italian regions.<sup>15</sup> In Sicily many young people feared to be forced to leave their village, “the only community to which they felt they belonged” (Rovinello, 2013, p. 498). Consequently, most of them fled to the Sicilian countryside to avoid the draft. Figure 1 illustrates the draft evasion phenomenon characterising Sicily between 1861 and early 1863.

The central government responded to draft resistance in Sicily with an intense campaign of military repression in Central and Western Sicily. From the end of June to mid-November of 1863 the military moved across Sicily to capture draft evaders and other criminals, under the command of General Giuseppe Govone. Troops, coming from different parts of Sicily, were first concentrated in the Caltanissetta and Agrigento provinces, deemed by the central government to be those most troubled by lawbreakers. As shown in Figure A1, in mid-August Govone’s soldiers moved west into the province of Trapani and in late September they entered Palermo, in the north of the island.<sup>16</sup>

---

<sup>14</sup>Cf. Duggan (2007), p. 289.

<sup>15</sup>A defining feature of the new national army was regiments made up of soldiers from diverse parts of the kingdom. ‘Mixing up’ the country’s population in the barracks was a deliberate policy by the central government, intended to unite Italians around a common national cause and identity (Duggan, 2007; Basile et al., 2022). However, this decision largely failed to make the soldiers feel more ‘Italian’ as a result. Outside their garrisons the soldiers had limited contact with the local population and did not learn new customs, whereas in the barracks they used to spend most of the time with other soldiers from their home region (Duggan, 2007, p. 288).

<sup>16</sup>This information mainly comes from military reports, letters and notes written by General Govone. See Sections 3.2 and 4.3 for more details.

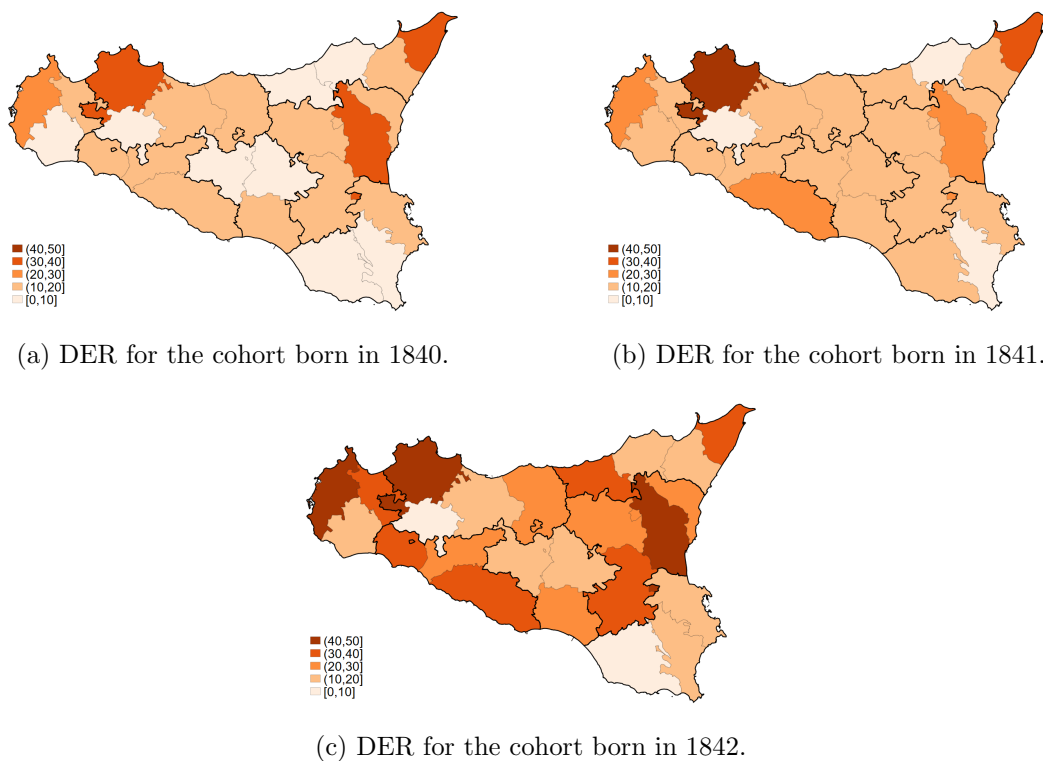


Figure 1: Draft evasion rate (DER, in p.p.) at the department (*circondario*) level for the Sicilian cohorts called up before the 1863 repression campaign. The black lines represent province boundaries. *Source*: Ministero della Guerra (1864).

Many sources describe the brutal tactics of this campaign. Military cordons were placed around towns and villages and frequent night raids carried out, with troops shooting anyone trying to escape (Davis, 1988, p. 184). To force local communities to reveal the hiding places of wanted men, the troops cut off towns' water supplies, persecuted the families of draft evaders and put entire towns under siege (Colajanni, 1900, p. 53). Other atrocities are also described by Riall (1998, p. 171):

“[...] in Petralia Soprana troops trying to burn one ‘rebel’ out of his house succeeded in killing almost his entire family; in Gangi they seized the pregnant wife of a wanted man, causing her to miscarry the child. The most damaging allegations concerned the torturing of a severely handicapped man named Antonio Cappello, on suspicion of simulating

muteness in order to avoid enlisting.”

Such barbaric measures were employed across multiple villages, even though the searches for conscripts were often based on inaccurate information.<sup>17</sup> Nevertheless, the ruling elite and the army leaders defended the use of these measures to capture ‘trouble-makers’, “to ensure the success of liberal policy [and] encourage... the honest people of Sicily” (Riall, 1998, p. 161). In the face of declining popular support for government, Govone’s troops increased the arbitrary use of force and by late August, the imposition of military cordons, the baseless arrests and indiscriminate penalties on entire towns and villages became the accepted norm for this campaign.

The search for draft evaders reached Eastern Sicily in mid-November 1863 and continued until the following spring under the supervision of General Giuseppe Medici. The measures employed by Medici’s troops in the three provinces of Messina, Catania and Noto were less severe than those used by Govone’s soldiers. Given the reported absence of resentment in the local population towards Medici’s military operations,<sup>18</sup> the focus of this work is entirely on the consequences of the military campaign in Central and Western Sicily in 1863.

### **2.3 Conscription and the Rise of the Mafia**

Conscription played an important role in increasing criminal activity in post-unification Sicily, for two reasons. First, the draft became a recruitment tool for brigand bands, whose ranks expanded considerably as more and more young men took to the hills after the call-up (Carraro, 1976, p. 95). Banditry posed a serious threat to the social order: in 1862, for instance, up to 142 crimes, including murders, thefts and highway robberies, were recorded in a single month in the province of

---

<sup>17</sup>As revealed by the Minister of War Della Rovere during the parliamentary debate held on December 5, 1863, out of the roughly 22,000 draft evaders and deserters reported on the conscription lists at the beginning of that year, about 8,000 of them were found to be female, dead or children by the end of Govone’s campaign (Riall, 1998, p. 173).

<sup>18</sup>“[Medici’s] troops behaved in such a way that no-one complained – they earned everyone’s goodwill and esteem. From this it can be seen that General Medici’s orders were given so as to not allow any power abuse from his soldiers” (*Il Precursore* newspaper, December 3, 1863, p. 3).

Palermo, while by May 1864 bandits were seen making incursions even into the city of Palermo (Davis, 1988, p. 183). Second, the 1863 military campaign led by General Govone to round up the draft evaders was, in the words of Alatri (1954, p. 76),

“the military expedition that left the greatest residues of resentment and hatred in the Sicilian population, due to the hateful methods of repression that were put in place [and] Govone’s declarations made before the Chamber of Deputies in defense of his actions.”<sup>19</sup>

As a consequence, state legitimacy and trust in the central government reached a new low in repressed towns. According to Dickie (2004, p. 57), “it was cases like [Antonio Cappello’s torture in 1863] that completely robbed the state of its credibility, and made many Sicilians very reluctant to trust it on any matter.” The sense of detachment from public institutions following Govone’s campaign was so strong that it lasted at least until the 1880s.<sup>20</sup>

The combination of the increase in brigandage and the decline in state legitimacy unintentionally cleared the way for the emergence of the Sicilian mafia.<sup>21</sup> The fear of draft evaders and brigands hiding in the Sicilian countryside and the

---

<sup>19</sup>During the parliamentary debate to which Alatri refers, held in December 1863 at the end of the campaign, Govone openly stated that “Sicilians were not yet ready for liberal government, as they had not yet developed beyond barbarism” (Riall, 1998, pp. 176-177).

<sup>20</sup>Sidney Sonnino, an Italian MP who personally ran an investigation on the economic conditions of Sicily together with Leopoldo Franchetti in 1876, summarises this concept during a speech before the Chamber of Deputies on March 30, 1881, as follows: “The vast majority of the population [...] feels entirely cut off from our institutions. People see themselves subjected to the state and forced to serve it with their blood and their money, but they do not feel that they are [an] organic part of it.”

<sup>21</sup>Gambetta (1993) claims that the lack of *autonomous* providers of protection in the Southern mainland, where the problems of brigandage and mistrust of the state were both tangible after the Italian unification, may have been the reason for the absence of mafia-type organisations spreading there too. Referring to the region of Calabria, Italy’s southern ‘toe’, Arlacchi (1983, p. 155) adds: “Where the violent and rebellious elsewhere became bandits and *mafiosi*, in the Crotonese area a significant number became paid hirelings of the barons... The private police of the latifundia remained fundamentally different from the mafia. They were, and always remained, dependent employees, mere executors of orders from above, without their own authority or even delegation of power.” In the same vein, Del Monte and Pennacchio (2012) find a negative correlation between the intensity of brigandage and the level of organised crime in continental Southern Italy after 1860.



enforcement officials searching for draft evaders generated widespread demand for private protection. At the same time, *mafiosi* managed to “leverage people’s discontent with the arbitrary and oppressive acts of the police state, and [they] gained the Sicilians’ solidarity” (Carraro, 1976, p. 113). This legitimisation process was crucial for the mafia to meet the stronger demand for security coming from a wider customer base. While the upper classes turned to the mafia for protection from bandits for themselves and their property, many peasant families also relied on the local *mafiosi* to guard their young family members from being conscripted, as they faced the risk of losing a key source of labour and income for several years.<sup>22</sup> As such, the newly formed Italian government unconsciously provided the protection industry with more opportunities for growth.

From this point onwards, the Sicilian mafia, by having “more and more people asking for its help rather than for legitimate enforcement,” accrued social legitimacy and power (Russo, 1964, p. 34). Increased legitimacy gave the mafia the ability to broaden its portfolio of services. As different segments of Sicilian society began to go to mafia bosses with other problems to solve, the mafia became involved in dispute resolution, cartel enforcement, and electoral canvassing (Franchetti, 1877; Lorenzoni, 1910; Hess, 1973). This process led to a “democratisation of violence,” as people from all social classes entered the protection industry (Franchetti, 1877).

By the mid-1870s the existence of the Sicilian mafia was *de facto* acknowledged both within and outside Sicily (Dickie, 2004, pp. 62-63). A key determinant of its emergence in that period were the actions of the Italian government, as explained by Colajanni (1900, pp. 63-64) when describing the condition of the island in the aftermath of Italy’s unification:

---

<sup>22</sup>“While the rich man could buy himself out of the draft [by paying a fee], the poor had to see his son leaving his home at the age of 20, just when he used to represent a capital [for his family]. This was something horrific in the eyes of the poor, who, being induced by the mafia, ended up preferring to see his son turning into a bandit, hidden in the local countryside, rather than into a soldier, far away [from home]” (Alongi, 1977, pp. 29-30).

“The action of the Italian government strengthened, instead of destroying, [...] the spirit that generated the mafia, i.e. the systematic mistrust of civil authority. The same Italian government acted in order to directly promote the spread of the mafia.”

## 3 Data

I exploit a novel historical dataset combining information on 350 Sicilian towns, organised into twenty-four departments (*circondari*) and seven provinces. The dataset includes the two main variables of interest, i.e. mafia presence and exposure to government repression in 1863, and a large set of historical information at the town level. Summary statistics are reported in Table A1.

### 3.1 Mafia Presence

The primary data on mafia presence come from the Borsani-Bonfadini parliamentary enquiry conducted between 1874 and 1876. The enquiry assessed the social and economic conditions of Sicily and the status of its public services. This enquiry, built upon more than 1,000 interviews with several local authorities and reported in Carbone and Grispo (1969), constitutes the earliest and most comprehensive collection of information on the conditions of public security at the town level in post-unification Sicily.

All province prefects in Sicily were issued a questionnaire on how the mafia operated in their own province, the social classes more likely to fall under the mafia’s influence, the main mafia-infiltrated towns and the names of the most dangerous *mafiosi* in their province. From the prefects’ answers I construct a dummy variable at the town level, *Mafia1875*, equal to 1 if a prefect recognized the presence of mafia in a given town, and 0 otherwise.<sup>23</sup> I identify 43 towns in the main sample (12.3% of the total) for which *Mafia1875* equals 1.

---

<sup>23</sup>This measure may suffer from under-reporting, as prefects might have been: unaware of the prevalence of mafia in their territories, fearful of retaliation or contiguity with the mafia, or they

Alternatively, Cutrera (1900) provides information on the early prevalence of mafia in Sicily and assesses the intensity of mafia activity in 287 Sicilian towns in 1900. He rates mafia intensity on a scale from none, to low, intermediate, and high.<sup>24</sup> I follow Buonanno et al. (2015) and code the level of mafia intensity from Cutrera (1900) using an ordinal variable, *Mafia1900*, ranging from 0 (no mafia) to 3 (major mafia presence). I take *Mafia1875* as my preferred dependent variable to analyse the effect of the 1863 repression campaign on the initial spread of the Sicilian mafia, and I use *Mafia1900* only to examine the persistent effects of exposure to government repression. Figures 2 and 3 show the geographical distribution of mafia activity, respectively, in 1875 and 1900. The Parliamentary Commission’s data suggest greater mafia presence in Western Sicily, especially in the Trapani province and in the areas near Agrigento and Palermo. A broadly similar pattern appears about twenty-five years later, according to Cutrera’s data.

### 3.2 Government Repression

The main explanatory variable, *Repr1863*, is a dummy defined at the town level, equal to 1 if the town was hit by repressive measures in 1863, and 0 otherwise. To identify the towns that experienced military violence, I mainly rely on very detailed information from a large number of personal notes, letters and military reports belonging to Giuseppe Govone, the army general in charge of the military

---

might not have had a clear understanding of what the term *mafia* meant, in which case they could not offer a valid assessment of this criminal phenomenon. Nevertheless, Dimico et al. (2017) indirectly suggest that potential under-reporting should not represent cause for concern in the estimation process. Whilst these scholars’ arguments apply to a different parliamentary enquiry, the Damiani-Jacini one, conducted in Sicily between 1881 and 1885, this public investigation was structured very similarly to the Borsani-Bonfadini enquiry.

<sup>24</sup>Antonino Cutrera was a police inspector in Sicily who used his own observations to document mafia activity across towns in 1900. Cutrera’s study is a key reference in previous studies on the Sicilian mafia’s origins (Gambetta, 1993; Buonanno et al., 2015; Dimico et al., 2017; Acemoglu et al., 2020) and is considered a reliable source by prominent mafia experts (Gambetta, 1993; Lupo, 1996). However, this source is from a period much later than Italy’s unification and it could be more informative about the Sicilian mafia’s evolution than about its emergence (Dimico et al., 2017, p. 1086).

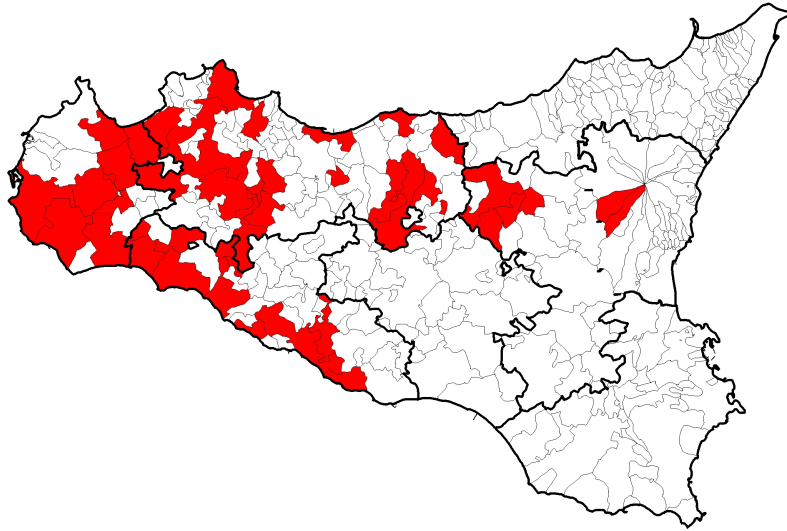


Figure 2: Mafia towns (in red) and non-mafia towns (in white) in Sicily in 1875. The black lines represent province boundaries. *Source:* Borsani-Bonfadini parliamentary enquiry, in Carbone and Grispo (1969).

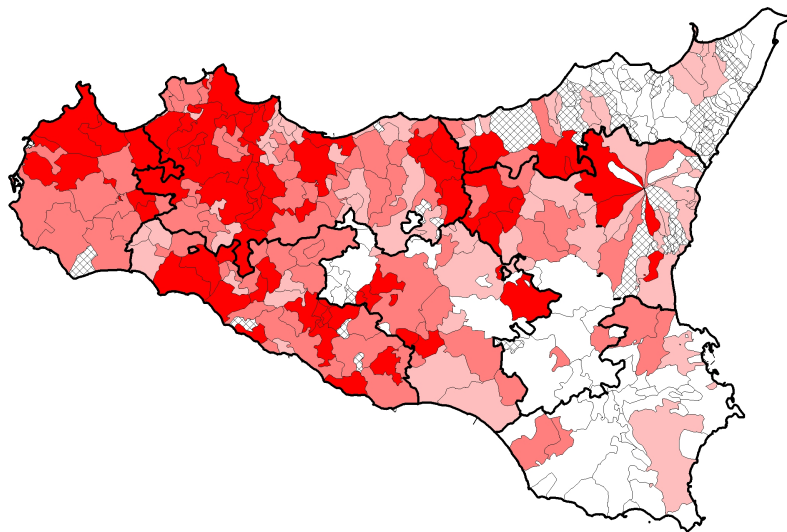


Figure 3: Mafia intensity at the town level in Sicily in 1900. Darker reds indicate more intense mafia activity, while towns with no mafia activity are reported in white (crosshatched towns indicate missing values). The black lines represent province boundaries. *Source:* Cutrera (1900).

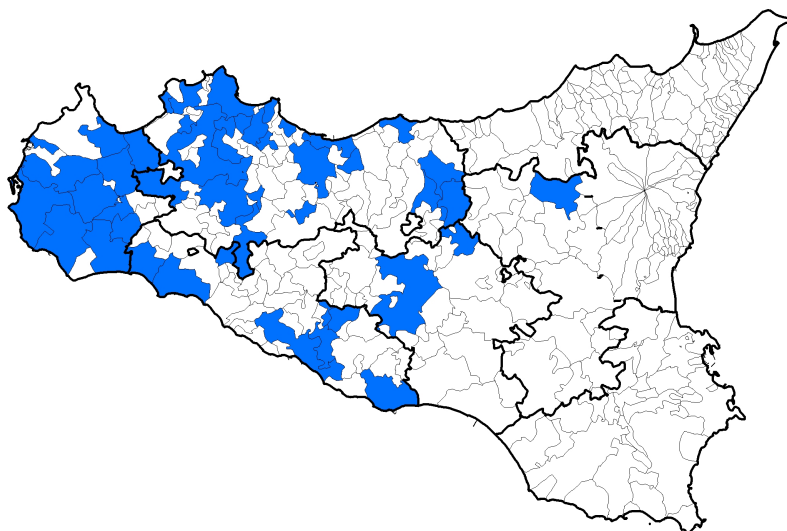


Figure 4: Towns hit by the government repression campaign in Sicily in 1863 (in blue). The black lines represent province boundaries. *Sources:* Govone Archive, *Giornale di Sicilia* (1863), *Il Precursore* (1863), *La Forbice* (1863), Maurici (1915), Brancato (1956) and Riall (1998).

operations in Sicily in 1863. These documents are primarily preserved in the Govone Archive in the National Museum of the Italian Risorgimento. I identify 76 repression episodes from these digitised primary sources, for which I know the exact geographical location, the event date and the type of offence committed.<sup>25</sup> Forty-three towns in the main sample have the value 1 for the variable *Repr1863* (12.3% of the total).<sup>26</sup> To reduce measurement error in the treatment variable, I also cross-check Govone’s papers with the information reported in three Sicilian daily newspapers providing local news coverage from June to December 1863: *Il Giornale Ufficiale di Sicilia* (pro-government), *Il Precursore* (anti-government) and *La Forbice* (anti-government).<sup>27</sup> Finally, I also rely on several historical sources covering the events in post-unification Sicily (Maurici, 1915; Brancato, 1956; Riall, 1998) to validate my key explanatory variable. Figure 4 shows the map of towns victimised by the 1863 repressive measures. The hardest-hit provinces (Agrigento, Trapani and Palermo) are all located in Western Sicily, which is consistent with the

military strategy chosen by the Minister of the Interior, Ubaldino Peruzzi, and the Minister of War, Alessandro Della Rovere, in early June 1863.<sup>28</sup>

### 3.3 Control Variables

To minimise omitted variables bias and improve the estimates' precision, I control for an extensive set of socio-economic and geographical variables at the town level, which are either correlated with exposure to government repression, level of mafia activity, or both.

The first set of control variables includes the drivers of the Sicilian mafia's early spread identified in the literature: citrus growing, sulphur presence, land concentration, and the suppression of the socialist movement known as Peasant Fasci. I proxy citrus production by the soil suitability index for the cultivation of citrus fruits computed by Buonanno et al. (2015), which builds upon data from the FAO-GAEZ project. I also include land suitability for cereals and olives, the two other most widespread crops in Sicily at that time, using the same data source. For sulphur, I use Parodi (1873) to construct a dummy variable that equals 1 for towns that produced sulphur between 1868 and 1870.<sup>29</sup> To account for different degrees of land concentration, I exploit information from the Damiani parliamentary

---

<sup>25</sup>To recognise a repression episode, I search for the following keywords in the above sources: *lagnanza* (complaint), *stato d'assedio* (town under siege), *arresto* (arrest) and *cordone* (military cordon).

<sup>26</sup>As mentioned in Section 2, the repressive measures put in place by the army in Sicily ranged from random arrests and detentions, to harassment, torture and killing. Although in some towns there were multiple types of offences against civilians, I choose not to exploit the information on the intensive margin of military repression, since gauging the severity of different offences would be subjective.

<sup>27</sup>The main reason for consulting different news outlets is that anti-government newspapers provided more detailed coverage of each stage of the repression campaign than pro-government ones.

<sup>28</sup>Cf. the letter from Peruzzi to Govone, June 1863, in Govone Archive, bundle n. 5, folder n. 5, n. 15.

<sup>29</sup>Focusing on the information on the extensive margin of sulphur availability rather than on the intensive margin helps reduce the concern about outliers possibly driving the results associated with this variable, a common issue in the resource curse literature (Buonanno et al., 2015).

enquiry (1885).<sup>30</sup> I index land concentration using an ordinal variable, ranging from 0 (low land concentration) to 2 (high land concentration). In the persistence analysis described in Section 6, where I focus on the effects on *Mafia*1900, I include a dummy variable for the presence of the leftist Peasant Fasci in the period 1891-1894 using data from Renda (1977), in the spirit of Acemoglu et al. (2020). Finally, to control for local criminal activity before the 1863 campaign, I use information on town-level banditry presence between 1820 and 1849 collected by Fiume (1984).<sup>31</sup>

The second set of covariates accounts for socio-economic and geographic characteristics. Following Buonanno et al. (2015) and Dimico et al. (2017), I include population density as a proxy for income per capita and urbanisation, using data from the 1861 Italian census and information on municipal surface area from Direzione Generale della Statistica (1867). I control for distance from the closest non-seasonal river and the presence of underground water basins, using information from OpenStreetMap<sup>32</sup> and the Sicilian Waters Observatory. Both factors are likely to have conditioned the demand for protection of produce, given the mafia's influence over the irrigation process through its control of water supply, as highlighted by Sylos Labini (2003). Data on average altitude and the altitude of the town centre are collected from the 1929 Agrarian Cadastre and Direzione Generale della Statistica (1886*b*). To control for the degree of remoteness, I use information on the postal distances to the regional capital (Palermo), to the respective department capital and to the closest commercial port from Lo Jacono (1856).<sup>33</sup> Finally, distance from the closest post road in 1856 is calculated from both the geo-referencing of a historical map of Sicily's road network (Smyth, 1826) and the complete set of post roads in

---

<sup>30</sup>For other uses of this source, see for instance Colaizzi (1895), Hess (1973), Lupo (1996), Bandiera (2003), Buonanno et al. (2015) and Dimico et al. (2017).

<sup>31</sup>See Fiume (1984), Fig. 11.

<sup>32</sup>The non-seasonal rivers network in Sicily in shapefile format is available at: <http://download.geofabrik.de/europe/italy/isole-latest-free.shp.zip> (last access: November 10, 2022).

<sup>33</sup>The main commercial Sicilian ports were Palermo, Catania, Messina, Siracusa, Porto Empedocle, Mazara del Vallo and Trapani.

1856 (Lo Jacono, 1856).<sup>34</sup>

Lastly, I use an extensive range of measures to capture heterogeneity in local state presence. As a proxy for both state reach into local communities and human capital endowment, I collect information on the number of primary schools from Direzione Generale della Statistica (1865), which refers to the school year 1862-1863. Importantly, this variable allows me to disentangle the effect of schooling as a concurrent nation builder from the effect of government repression on mafia presence, prior to the historical event I discuss in this study. To capture differentials in civic-ness, trust in institutions, and social identification with the Piedmontese rulers, I use two measures: electoral turnout in 1861, with data collected and digitised from the historical archive of the Italian Chamber of Deputies, and a dummy variable (*Garibaldini* 1860) constructed using information on the town of origin of about 1,000 Sicilians who voluntarily joined General Garibaldi during the 1860 unification campaign.<sup>35</sup> Following the mechanism proposed by Lecce et al. (2022), the presence of *garibaldini* in a town may reflect a *positive* inclination toward the central government by the Sicilian population living there, before the implementation of the widely unpopular reforms discussed earlier. Finally, to account for the presence of the Catholic Church, I also compute a dummy variable that takes the value of 1 if the town was one of the thirteen episcopal sees of the Catholic Church in Sicily in 1839, using data from De Sanctis (1840).<sup>36</sup>

Table A2 reports mean values of the control variables described above for repressed and non-repressed towns. Treatment and control groups are unbalanced along several observable characteristics, especially those related to local state presence, thus in subsequent calculations I put more emphasis on the specifications

---

<sup>34</sup>Given the high quality of Smyth (1826)'s map, the post road distance variable that I construct represents an improvement over other proxies for distance from the post road network used in the literature, all based on Cary (1799).

<sup>35</sup>For this information I employ data scraping from the website of the State Archive of Turin, in which individual-level data on more than 35,000 *garibaldini* are stored.

<sup>36</sup>To construct this variable I consider towns that had a bishopric or an archbishopric in 1839 (see De Sanctis, 1840, p. 32).



that include the entire set of possible confounders. In Table A3 I also show maximum likelihood results from a logistic regression in which I use *Repr1863* as dependent variable. Local indicators of state presence strongly negatively predict the probability of being exposed to government repression. This finding is consistent with the objective of the repression campaign, i.e. the search for draft evaders, as it is likely that they were more present in places characterised by a lower pre-existing level of state legitimacy.

## 4 Results

In this section I report results from ordinary least squares (OLS), conditional logit, and 2SLS regressions of mafia presence in 1875 on exposure to government repression in 1863. My baseline equation is:

$$Mafia1875_{ip} = \alpha + \beta_1 Repr1863_{ip} + x_{ip}'\beta_2 + \mu_p + \epsilon_{ip}, \quad (1)$$

where *Mafia1875<sub>ip</sub>* represents a dummy variable for the presence of mafia in 1875 in town *i* and province *p*, while *Repr1863<sub>ip</sub>* denotes a dummy variable equal to 1 if government repression occurred in 1863 in town *i* and province *p*. *x<sub>ip</sub>* is a vector of covariates that may affect the probability of mafia presence, as described in Section 3.3.  $\mu_p$  indicates province fixed effects and  $\epsilon_{ip}$  is the error term including all omitted variables. Fixed effects and distance controls ensure that I compare local differences in mafia presence. I expect  $\beta_1$  to be positive and significantly different from zero, suggesting that, conditional on observables, being violently exposed to a coercive nation-building policy in 1863 is associated with a higher probability of the Sicilian mafia spreading in the municipal territory in the medium term. I present estimates of standard errors clustered at the department level, to account for the fact that, starting in 1861, the Sicilian administrative, judicial and law enforcement systems were all organised at this administrative level. In Section 4.2 I also report Conley

(1999)’s spatially-corrected standard errors for robustness.

## 4.1 OLS Results

Table 1 reports OLS estimates of Equation (1) on the sample of 350 towns. The different columns in this table gradually increase the number of covariates. Column 1 shows a raw correlation between mafia presence and government repression of 0.443, which is significant at the 1% level. Adding province fixed effects in Column 2 decreases the magnitude slightly but does not affect its significance level. The reason for introducing province fixed effects is twofold. On the one hand, as explained in Section 3.1, the variable  $Mafia1875_{ip}$  depends on the information provided by each of the prefects operating in Sicily in 1875, who were responsible for maintaining public order in their respective province. Including province fixed effects in the linear probability model described in Equation (1) would therefore allow me to identify the effect of government repression within small areas characterised by a fairly homogeneous level of institutional, or ‘prefect’, quality. On the other hand, province fixed effects base inference on within-province variation in mafia presence and thus prevent the results from being driven by broader differences between Eastern and Western Sicily.

Columns 3 to 5 add, in order, well-known determinants of the early prevalence of the Sicilian mafia (Fiume, 1984; Buonanno et al., 2015; Dimico et al., 2017), an extensive battery of exogenous geographical controls, and a set of pre-1863 state presence confounders. The coefficient of interest remains always positive and highly significant, regardless of controls. The inclusion of land concentration controls in Column 6 does not affect the sign and significance of  $\beta_1$ , but substantially reduces the sample size. Therefore, the specification in Column 5 remains my preferred one, as it allows for larger cross-town variation.

To gain a sense of the potential impact of unobservables, I rely on Altonji et al. (2005) and Oster (2019), who propose calculating the degree of selection on

Table 1: Baseline Estimates

Dependent variable:	Mafia 1875					
	(1)	(2)	(3)	(4)	(5)	(6)
Repression 1863	0.443*** (0.098)	0.338*** (0.075)	0.344*** (0.076)	0.314*** (0.076)	0.274*** (0.089)	0.252** (0.091)
Land suitability for citrus			-0.005 (0.004)	0.003 (0.006)	0.004 (0.006)	0.003 (0.006)
Land suitability for olive			0.003 (0.003)	-0.000 (0.003)	-0.001 (0.003)	-0.001 (0.003)
Land suitability for cereals			0.001 (0.004)	-0.001 (0.005)	-0.002 (0.005)	-0.002 (0.005)
Active sulphur cave 1868-1870 (dummy)			-0.112* (0.063)	-0.108* (0.052)	-0.126** (0.057)	-0.168** (0.071)
Banditry presence 1820-1849 (dummy)			0.079* (0.045)	0.088* (0.048)	0.072 (0.052)	0.079 (0.066)
Ln population density 1861					0.011 (0.010)	0.016 (0.014)
Elementary schools 1862 per 1,000 people					-0.033 (0.028)	-0.037 (0.032)
Electoral turnout 1861					0.060 (0.129)	0.060 (0.156)
Episcopal see 1839 (dummy)					0.251** (0.121)	0.265* (0.129)
Garibaldi presence 1860 (dummy)					0.058* (0.033)	0.060* (0.031)
Medium land concentration (dummy)						-0.012 (0.039)
High land concentration (dummy)						0.078 (0.058)
Province FE	No	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes	Yes
N	350	350	350	350	350	294
No. of clusters	24	24	24	24	24	24
P-value of wild cluster bootstrap test	0.000	0.000	0.002	0.003	0.001	0.000
Mean of dependent variable	0.123	0.123	0.123	0.123	0.123	0.133
R <sup>2</sup>	0.196	0.275	0.293	0.333	0.365	0.399
Oster (2019)'s $\delta$		1.569	1.586	1.579	1.419	1.309

*Notes:* OLS estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1875 at the town level. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. In column (6) the dummy *Low land concentration* is omitted from the model. Standard errors clustered by department are reported in parentheses. The p-value of the wild cluster bootstrap test of the hypothesis  $Repr_{1863} = 0$  is reported (bootstrap replications = 999). Oster (2019)'s relative degree of selection  $\delta$  is estimated using  $R_{max} = 1.3R^2$ . \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

unobservables relative to observables that would be necessary to explain away the main result. A very high relative degree of selection would suggest that unobservables would need to have disproportionate relevance with respect to observables in order to produce a treatment effect equal to 0. Oster (2019) refers to this degree via the parameter  $\delta$ , which I report in Table 1.<sup>37</sup> Since in Columns 2 to 6 the estimated  $\delta$  is always greater than 1, these results suggest that unobservables are unlikely to be sufficiently important to eliminate the observed effects.

## 4.2 Robustness

The OLS estimates in Table 1 clearly indicate that exposure to government repression and mafia presence correlate positively. To bolster confidence in my baseline results, I run a battery of robustness checks.

I first replicate the analysis of Table 1, including department (*circondario*) fixed effects instead of province fixed effects. After comparing towns within even smaller administrative areas, the main results still hold (Table A4). Then I estimate Conley (1999)’s standard errors to account for spatial correlation, using the initial empirical setting with province fixed effects. I alternatively use a cut-off distance of 15 km, 30 km, or 0.4 degree for latitude and 0.75 degree for longitude, which represent roughly half of the height and width of the Sicilian island (Acemoglu et al., 2020). The correction for spatial correlation delivers standard errors that are similar or slightly more conservative than the department-clustered errors used in the main analysis (Table A5).

Since *Mafia1875* is a binary outcome variable, I also estimate a conditional logistic model, controlling for unobserved heterogeneity among different departments

---

<sup>37</sup>A key input into the estimation of  $\delta$  is a theoretical maximum R-squared from a hypothetical regression of the outcome on the treatment and both observed and unobserved covariates. Oster (2019) suggests 1.3 times the R-squared from the baseline specification (e.g.  $1.3 \times 0.365 = 0.475$  as per the specification in Column 5), especially for empirical settings in which the outcome cannot be fully explained even if the full control set is employed, for example because of measurement error affecting the dependent variable (as in this case, see Footnote 23). In my analysis I follow Oster’s suggestion and consider  $R_{max} = 1.3R^2$ .

and clustering standard errors at the same administrative level. Table A6 reports maximum likelihood estimates of *Mafia*1875 on *Repr*1863, after progressively adding control variables as in Table 1. This specification reduces the sample size from 350 to 169 towns with the inclusion of department fixed effects.<sup>38</sup> The preferred specification (Table A6, Column 5) shows that the odds of being a mafia-ridden town in 1875 are about  $\exp(1.548) = 4.7$  times as large for towns that were exposed to government repression in 1863 relative to towns that were not. The results in Table A6 suggest once again that, within each department, variations in government repression are positively and significantly associated with variations in mafia presence.

Lastly, to prevent potential biases due to the non-random assignment of treatment, in the left panel of Table A7 I use nearest neighbour matching estimation on my preferred specification. On the right panel I also employ matching estimation by geographical location, based on latitude and longitude only, to exploit variations in government repression across direct neighbours. This exercise allows for the comparison of towns no more than a few kilometres apart, where one was subject to the repressive measures in 1863 while the others were not.<sup>39</sup> The impact of repression identified when matching one treated town with the two nearest untreated towns (Table A7, Column 5) is significant and of a very similar magnitude to the baseline estimates.<sup>40</sup> Overall, the OLS estimates of the Average Treatment Effect for the

---

<sup>38</sup>Each of the towns located in the twelve departments lacking mafia presence in 1875 does not contribute to the estimation of  $\beta_1$ , so it is automatically dropped from the analysis. The substantial reduction in sample size entailed by the estimation of this discrete choice model represents the main reason for favouring the linear probability model, despite the latter being less efficient.

<sup>39</sup>To estimate the Average Treatment Effect for the Treated (ATET), I assume that, on average, towns in the treatment group would perform as much as those in the control group if they were not treated, conditional on observables. More formally, I assume that  $\mathbb{E}(Mafia1875^0|x, Repr1863 = 1) = \mathbb{E}(Mafia1875^0|x, Repr1863 = 0)$ .

<sup>40</sup>This result is consistent with Słoczyński (2022), who shows that the Average Treatment Effect (ATE) of a binary treatment variable in a linear model with additive effects is, when estimated with OLS and under certain conditions, a convex combination of the Average Treatment Effect for the Treated (ATET) and for the Untreated (ATEU). Słoczyński (2022) also finds that the weights placed by OLS on the ATET and ATEU are inversely related to the proportion of observations in each group, so that e.g. the less units are treated, the more the OLS-estimated ATE is biased towards the ATET, and viceversa.

Treated (ATET) in Table A7 provide compelling evidence that my findings are not driven by unobserved heterogeneity at the local level.

### 4.3 IV Strategy

Despite the rich set of controls included in Equation (1), the high value of Oster (2019)'s estimated parameter  $\delta$  and the results from propensity score and geographical matching estimation in Table A7, the positive and significant effect of exposure to government repression on early mafia diffusion may still be prone to endogeneity bias. Two potential threats to identification could preclude the causal interpretation of the OLS estimates in Table 1: measurement error affecting the explanatory variable *Repr1863* and selection into treatment. About the former, the information included in Govone's documents, my main source to keep track of the military campaign, may be incomplete or inaccurate. In addition, the events occurring in some repressed towns may have gone unnoticed and did not appear in local newspapers.<sup>41</sup> Both issues would produce non-classical measurement error in the recording of exposure to government repression, leading to an attenuation bias in the OLS estimates (Bingley and Martinello, 2017). On selection, instead, the military may have targeted towns initially characterised by low crime rates, in order to exert less effort in the search for draft evaders. Alternatively, troops may have focused on areas with a high pre-existing level of criminal activity, where draft evaders would be more likely to locate. In either case, the resulting endogeneity of government repression would bias the OLS results. To address the concern with selection into treatment, I propose an instrumental variables framework based on the military plan of attack underlying the 1863 repression campaign, illustrated in General Govone's

---

<sup>41</sup>The purposeful omission of repression episodes, while being a common feature of pro-government newspapers, seems quite unlikely to have affected the entire spectrum of news outlets circulating across Sicily in 1863. In the same vein, the spreading of false information about the military campaign or the reporting of fake repression episodes in newspapers should not represent a matter of concern either, as anti-government newspapers used to adopt fact-checking practices whenever possible (cf. *Il Precursore*, September 29 and October 14).

papers and briefly summarised below.

Since the start of the expedition in late June, Govone's military strategy was based on speed and surprise. Govone ordered his military columns to move swiftly from one location to the next and also instructed them to cordon off the towns met on their way before dawn without notice.<sup>42</sup> The adoption of a 'hit-and-run' strategy was partly determined by the limited amount of military forces involved in the campaign and the consequent need to let soldiers return to their original provinces, in order not to compromise the maintenance of public order there. This issue became particularly relevant in early August, when Govone urged his troops to bring the military operations in the Agrigento province to completion as soon as possible, so as to let the military promptly return to the provinces of Trapani and Palermo.<sup>43</sup>

I leverage the information above on the military plan of attack sponsored by General Govone to generate exogenous variation in exposure to government repression, in four steps. First, I identify the set of Sicilian towns that had both military barracks prior to the 1863 campaign and represented main stops made by Govone's troops on their way. I thus obtain a list of 10 towns located in Central and Western Sicily, most of them being department capitals before the 1863 campaign.<sup>44</sup> Second, I construct least-cost paths connecting the selected garrison towns to each other, conditional on the chronological order in which such places were visited by the army (Figure A1).<sup>45</sup> I define these least-cost paths, or simply shortest routes,

---

<sup>42</sup>See Govone's instructions to the troops on June 28, 1863, in Govone Archive, bundle n. 5, folder n. 3, n. 30.

<sup>43</sup>The reason for this request was a worrying deterioration of public order in those two provinces, due to the insufficient military presence (cf. Govone Archive, bundle n. 5, folder n. 5, n. 21).

<sup>44</sup>The garrison towns here identified are, in chronological order of visit, Piazza Armerina, Terranova di Sicilia, Agrigento, Marsala, Trapani, Calatafimi, Palermo, Corleone, Cefalù and Polizzi Generosa.

<sup>45</sup>During the 1863 campaign General Govone relied on the support of different platoon leaders, who were responsible for coordinating the military operations in the limited geographical areas under their own specific authority. To construct the optimal route network, I therefore identify the least-cost path connecting a garrison town to the one that came immediately next in the itinerary of each platoon leader. This implies that a given garrison town can represent the starting or ending

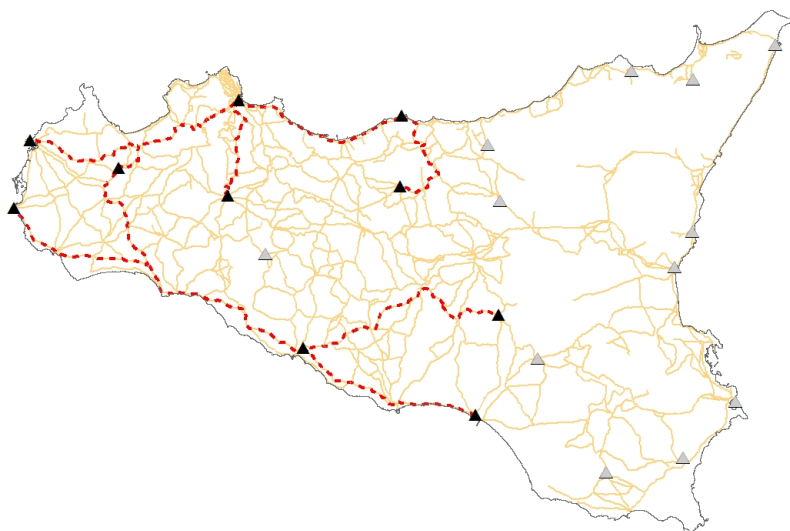


Figure 5: 1826 road network (light brown solid line), garrison towns met by Govone’s troops (black triangles) and optimal route network (red dotted line) connecting them. Grey triangles are used to represent the garrison towns that were not involved in the 1863 repression campaign. *Sources:* Smyth (1826) and Govone (1863).

using the pre-existing road network map drawn by Smyth (1826), which I digitised in GIS.<sup>46</sup> Figure 5 shows the optimal route network. Third, I calculate the distance of each town in the dataset to these shortest paths, and I use it as an instrumental variable for the exposure to government repression. Fourth, I remove the selected garrison towns from the sample, to minimise selection into treatment.

This procedure solves a number of problems of endogeneity associated with the treatment variable. First of all, the constructed network of least-cost paths acts as a proxy for the route that *could* have been taken by the troops, not the *actually* taken route, which would be most likely endogenous. Using the optimal route network in Figure 5 allows me to generate a quasi-random group of towns that were more likely to suffer from government repression only because of their proximity to the shortest route between two garrison towns, in the spirit of the “inconsequential point of distinct least-cost paths.”

<sup>46</sup>It is worth noting that moving around by train was not an option for Govone’s troops. In fact, railways were almost nonexistent in Sicily in 1863. By the end of that year there existed only one railway line running between Palermo and Bagheria, about 13 kilometers long.



places” approach in Redding and Turner (2015). Finally, the shortest route network also allows me to fairly account for the ‘hit-and-run’ nature of the military operations conducted in Central and Western Sicily in 1863.

#### 4.4 2SLS Results

I test whether the proximity to the shortest route network is correlated with exposure to the government repression campaign of 1863 by estimating the following first-stage relationship:

$$Repr1863_{ip} = \lambda + \gamma_1 Distance_{ip} + x_{ip}'\gamma_2 + \mu_p + \nu_{ip}, \quad (2)$$

where  $Distance_{ip}$  is the natural logarithm of (one plus) the shortest distance of each town to the optimal route network. The vector of town-level covariates  $x_{ip}$  and the province fixed effects  $\mu_p$  are the same as in Equation (1). The use of a log-transformed version of the distance measure is motivated by the typical presence of small networks of towns in 19th-century Sicily, causing a non-linear relationship between physical distance and probability of social interaction between people living in different areas.<sup>47</sup>

Table 2 reports OLS and 2SLS estimates of Equation (1) on the restricted sample of 340 towns, obtained after removing the garrison towns used to construct the shortest route network. Panel A shows 2SLS estimates of Equation (1), Panel B reports the corresponding first stage, Panel C adds reduced-form specifications and Panel D presents OLS estimates of the coefficient  $\beta_1$  for comparison. Column 1 shows a positive and significant effect of exposure to government repression on the early spread of the Sicilian mafia. The estimated coefficient  $\hat{\beta}_1$  also does not significantly change in either magnitude or significance when more controls are progressively included (Columns 2 to 5). The preferred specification (Column 5) shows that exposure to the repression campaign increased the probability of early

---

<sup>47</sup>This is in line with the argument made by Lecce et al. (2022) for continental Southern Italy.

Table 2: IV Analysis

	(1)	(2)	(3)	(4)	(5)
Panel A: Second stage	Dependent variable: Mafia 1875				
Repression 1863	0.812*** (0.160)	0.851*** (0.217)	0.889*** (0.241)	0.809*** (0.216)	0.788*** (0.222)
Mean of dependent variable	0.106	0.106	0.106	0.106	0.106
Panel B: First stage	Dependent variable: Repression 1863				
ln(1 + Distance from least-cost path)	-0.104*** (0.016)	-0.096*** (0.025)	-0.092*** (0.025)	-0.100*** (0.026)	-0.100*** (0.026)
Kleibergen-Paap Wald F statistic	41.83	14.48	13.16	15.08	15.02
Panel C: Reduced form	Dependent variable: Mafia 1875				
ln(1 + Distance from least-cost path)	-0.084*** (0.017)	-0.081*** (0.023)	-0.082*** (0.026)	-0.081*** (0.022)	-0.079*** (0.023)
P-value of wild cluster bootstrap test	0.000	0.006	0.008	0.013	0.016
R <sup>2</sup>	0.143	0.210	0.233	0.273	0.301
Panel D: OLS Estimates	Dependent variable: Mafia 1875				
Repression 1863	0.379*** (0.108)	0.287*** (0.079)	0.284*** (0.083)	0.268*** (0.081)	0.245*** (0.089)
P-value of wild cluster bootstrap test	0.000	0.002	0.006	0.003	0.002
R <sup>2</sup>	0.143	0.236	0.256	0.297	0.316
Province FE	No	Yes	Yes	Yes	Yes
Main controls	No	No	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes
State presence controls	No	No	No	No	Yes
N	340	340	340	340	340
No. of clusters	24	24	24	24	24

*Notes:* 2SLS (Panel A) and OLS (Panel D) estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1875 at the town level. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. This regressor is instrumented by the natural logarithm of (one plus) the distance to the network of least-cost paths described in Section 4.3. The ten garrison towns used to construct this optimal route network are dropped from the sample. Panel B reports the results of the first-stage regression for the exposure to government repression, whereas Panel C includes estimates of the reduced-form regression of mafia presence on the excluded instrument. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. Standard errors clustered by department are reported in parentheses. The p-value of the wild cluster bootstrap test of the hypothesis  $Repr1863 = 0$  is reported (bootstrap replications = 999). \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

mafia presence by approximately 79 percentage points, compared to a mean of 10.6% (on the restricted sample).

To substantiate the validity of this instrumental variables strategy, I first report both the first-stage regressions for all IV estimations and the corresponding reduced-form specifications. For my second-stage estimates to be reliable, I need a sufficiently strong correlation between the instrument and the repression variable. This condition is confirmed by both the first-stage F-statistics shown in Panel B and the OLS estimates from the reduced-form regressions in Panel C. Next, I address possible violations of the exclusion restriction. This condition requires the effect of proximity to the shortest route network on mafia presence to work solely through the exposure to the 1863 repression campaign. Any direct effect of closeness to the optimal network on my outcome of interest would therefore be a threat to instrument validity. In Table 3 I provide evidence against this possibility by running three falsification tests on the restricted sample. In Panels A, B and C I report estimates of the relationship between the instrument and, respectively, a proxy for pre-existing criminal activity (banditry presence between 1820 and 1849) and two proxies for ex-ante attitudes towards the central government (the *garibaldini* binary variable and the electoral turnout in 1861). Irrespective of the specification being considered, proximity to the shortest route network is uncorrelated with all these pre-1863 characteristics, increasing my confidence in the exclusion restriction underlying this empirical approach.

Lastly, it is possible that differences in proximity to the optimal route network may partly pick up the effect of closeness to main roads, like commercial routes, which may be relevant for the early diffusion of the Sicilian mafia. Since the selected garrison towns being employed to identify the least-cost paths often represented major towns, this threat could be particularly relevant. However, in the most demanding specifications in Table 2 (Columns 4 and 5) I control for distance to the post road network, which includes the main commercial roads existing in Sicily

Table 3: Falsification Tests

	(1)	(2)	(3)	(4)
Panel A	Dependent variable: Banditry presence 1820-1849			
ln(1 + Distance from least-cost path)	-0.041 (0.026)	-0.035 (0.027)	-0.026 (0.031)	-0.021 (0.031)
P-value of wild cluster bootstrap test	0.160	0.320	0.468	0.588
R <sup>2</sup>	0.335	0.358	0.378	0.422
Mean of dependent variable	0.244	0.244	0.244	0.244
Panel B	Dependent variable: Garibaldi 1860			
ln(1 + Distance from least-cost path)	-0.061 (0.037)	-0.040 (0.034)	-0.020 (0.034)	-0.025 (0.031)
P-value of wild cluster bootstrap test	0.119	0.259	0.564	0.438
R <sup>2</sup>	0.087	0.134	0.161	0.227
Mean of dependent variable	0.438	0.438	0.438	0.438
Panel C	Dependent variable: Turnout 1861			
ln(1 + Distance from least-cost path)	-0.003 (0.011)	-0.002 (0.012)	-0.007 (0.011)	-0.007 (0.011)
P-value of wild cluster bootstrap test	0.827	0.834	0.584	0.542
R <sup>2</sup>	0.030	0.043	0.066	0.100
Mean of dependent variable	0.840	0.840	0.840	0.840
Province FE	Yes	Yes	Yes	Yes
Main controls	No	Yes	Yes	Yes
Geographic controls	No	No	Yes	Yes
State presence controls	No	No	No	Yes
N	340	340	340	340
No. of clusters	24	24	24	24

*Notes:* OLS estimates of the relationship between the natural logarithm of (one plus) the distance to the network of least-cost paths described in Section 4.3 and pre-repression variables. The ten garrison towns used to construct the optimal route network are dropped from the sample. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldi* in 1860. In Panel A the banditry dummy variable is removed from the set of controls and used as dependent variable. In Panel B the *garibaldi* dummy variable is removed from the set of controls and used as dependent variable. In Panel C the 1861 turnout variable is removed from the set of controls and used as dependent variable. Standard errors clustered by department are reported in parentheses. The p-value of the wild cluster bootstrap test of the hypothesis  $Repr1863 = 0$  is reported (bootstrap replications = 999). \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

in 1863. Controlling for this covariate does not affect the coefficient of interest. In any case, in the remainder of the analysis I put more emphasis on the results from specifications that account for this confounder.

Table 2 also shows that the 2SLS estimates are consistently larger than the corresponding OLS estimates. There are a number of reasons that could explain this finding. One explanation would be the presence of measurement error affecting the exposure to government repression. Since my treatment is binary, the potential measurement error associated with it is non-classical by construction, leading to an attenuation bias in OLS and an inflation bias in IV (Bingley and Martinello, 2017). Moreover, the OLS estimates may be biased downward due to omitted variables that predict exposure to the treatment but decrease the probability of mafia presence, or the exclusion restriction could be severely violated. Another potential explanation is the difference between the Local Average Treatment Effect (LATE) picked up by the IV estimates and the Average Treatment Effect (ATE) estimated through OLS and reported in Panel D.

The results in Tables 3 and A7 suggest that omitted variables and violations of the exclusion restriction are both unlikely to drive the difference between 2SLS and OLS estimates. Although measurement error in the treatment variable *Repr1863* cannot be neglected, it seems more likely that the reported gap in estimated effects could be due to the fact that my instrument focuses identification on compliers, i.e. towns that were exposed to government repression because of their proximity to a least-cost path. Therefore, towns in the approximate path of the armies could have displayed a much stronger effect of repression on subsequent mafia activity than the one exhibited by the average town in the sample.<sup>48</sup>

I conclude this section with three robustness checks, in which I assess different transformations of the distance variable used as excluded instrument. Table A8

---

<sup>48</sup>Though being applied to a different context, this reasoning is similar to the one presented by Tur-Prats and Valencia Caicedo (2020), who look at the effect of political violence on the routes used by rebel troops during the Spanish Civil War.

provides 2SLS estimates of the impact of exposure to government repression on the emergence of the Sicilian mafia, where treatment is instrumented by either the inverse hyperbolic sine transformation of the distance to the shortest route network (Panel A) or a dummy variable equal to 1 if a town falls within 5 kilometers of a least-cost path (and set to 0 otherwise) (Panel B). The former transformation reduces the sensitivity of my results to the specific constant value to add before applying the log transform, in order to avoid zero distances, while the latter takes into account possible inaccuracies in the historical road network that I use to compute the distance variable.<sup>49</sup> Both instruments deliver second-stage estimates that are very close to the main 2SLS results in Table 2, increasing my confidence in this IV strategy. Finally, in Table A9 I report Conley (1999)’s spatially-corrected standard errors for Equation (1), using again the log distance to the optimal route network to predict *Repr*1863. The newly estimated standard errors do not alter the statistical significance of the coefficient of interest  $\beta_1$ , regardless of the specification being considered.

## 5 Mechanisms

The empirical findings reported so far are all indicative of a strong and positive relationship between government repression and the early spread of the mafia in Sicily. I now explore two potential mechanisms through which violent exposure to a coercive nation-building policy may foster organised crime: increasing demand for mafia by the local population, and increasing mafia supply.

### 5.1 The Demand for Mafia

Although until the mid-1870s Sicily was the scene of intense disorders involving the Italian army,<sup>50</sup> Govone’s military campaign in Sicily in 1863 represented a

---

<sup>49</sup>The distance dummy method is in the same vein as the approach used by, for instance, Jedwab and Moradi (2016) and Fenske et al. (2021).

<sup>50</sup>Deserving of note were especially the large-scale military operations ordered by local public authorities in 1865 to persecute both real and perceived political enemies, and the military suppression of the *Seven and a Half Days Revolt* in Palermo in 1866. The former concentrated

turning point in the relationship between the recently-born state and its Sicilian periphery. Most importantly, historians generally agree that these operations severely compromised both state legitimacy and trust in the central government (Colajanni, 1900; Alatri, 1954; Dickie, 2004). Govone's military sweeps were often in disregard of civil liberties granted by law, and consciously so.<sup>51</sup> As a result, the local communities in Sicily soon began to reject the central government's right to rule and its position as the highest political authority. This phenomenon affected the protection industry in a meaningful way: by repressing local people, the central government increased incentives for many Sicilians to demand protection, but also lost its credibility as a trustworthy provider of protection. Such a transformation of the market for protection ultimately favoured the mafia, which emerged as a reliable protection supplier and consequently gained a broad social consensus.

To substantiate the above arguments, I empirically investigate the effect of government repression on the level of trust in the central government in Sicily after 1863, which I measure at the town level in two ways. First, I use voter turnout as a proxy for trust in institutions by the local elites (Guiso et al., 2000; Cannella et al., 2021).<sup>52</sup> I construct this variable by using newly digitised data from the parliamentary elections in 1861 (to conduct a placebo test), 1865 and 1867, collected from the historical archive of the Italian Chamber of Deputies. Second, I employ as a dependent variable the ratio of vaccinated children between 1880 and 1884 in the

---

across the province of Palermo and were characterised by high levels of military discipline and few reported cases of abuses (Riall, 1998, p. 189). The latter episode was much more violent and affected instead only the regional capital and its neighboring villages.

<sup>51</sup>Govone often acknowledged the illegitimacy of methods adopted by his troops, but was unconcerned about such a *modus operandi*, for the sake of public order: "relying on legal means will not be enough to stop [the draft evasion phenomenon] anytime soon" (cf. letter from Govone to Alfonso La Marmora, July 28, 1863, in Govone Archive, bundle n. 4, folder n. 8, n. 2d).

<sup>52</sup>After Italy's unification, all the Southern regions, including Sicily, were forced to adopt the Kingdom of Sardinia-Piedmont's electoral rule, which had very restrictive voter eligibility requirements in terms of gender, age, literacy level and income. According to this electoral law, only males older than 25, being able to earn at least 40 liras per year and with proven reading and writing skills, were eligible to vote (Accetturo et al., 2017, p. 229). The combination of these criteria reduced the total number of eligible voters in Sicily to roughly 38,000 in 1865, corresponding to approximately 1.6% of the total population on the island in 1861.

total population in 1881, sourced from Direzione Generale della Statistica (1886*b*), to capture the effect of exposure to the 1863 repression campaign on trust in the central government across all social classes.<sup>53</sup> I run a standard OLS regression to estimate the following cross-sectional relationship:

$$y_{ip} = \tau + \phi_1 Repr1863_{ip} + x_{ip}'\phi_2 + \mu_p + \xi_{ip}, \quad (3)$$

where  $y_{ip}$  is a proxy for trust in government,  $Repr1863_{ip}$  indicates exposure to government repression in 1863, and  $\mu_p$  denotes province fixed effects as before. When the dependent variable is electoral turnout, I add to the usual set of confounders the straight-line distance of each town to the respective polling station in each election year.<sup>54</sup>

Table 4 reports the OLS estimates of Equation (3). There is no statistically significant association between exposure to government repression in 1863 and pre-existing trust in institutions by the Sicilian electorate (Column 1). On the contrary, when I use voter turnout in 1867 as the dependent variable (Column 3), the estimated coefficient  $\hat{\phi}_1$  is negative and strongly significant, in line with the distrust-in-government mechanism proposed above.<sup>55</sup> More specifically, I find that experiencing government repression in 1863 results in a lower electoral turnout in

---

<sup>53</sup>The use of vaccine uptake to measure changes in trust in institutions is not new in the literature. In the late 1990s distrust in the British government was linked to hesitancy towards the measles-mumps-rubella (MMR) vaccine across the UK, according to Larson and Heymann (2010). More recently, a positive relationship between trust in government and rates of compliance with public health policies has been documented by Blair et al. (2017) during the 2014–15 Ebola epidemic in Liberia and by Bargain and Aminjonov (2020) and Durante et al. (2021) in the context of the COVID-19 pandemic. On the relationship between vaccines and trust, see also Lowes and Montero (2021) and Martinez-Bravo and Stegmann (2022).

<sup>54</sup>In 1861, and only in that year, a polling station was set up in each town, so this variable is not included in the specification in Table 4, Column 1.

<sup>55</sup>This result does not contradict the findings of Accetturo et al. (2017), who document an *increase* in electoral turnout in the provinces of continental Southern Italy where the central government introduced a form of martial law in August 1863, known as Pica Law, to repress the brigandage. The government’s fight against brigandage in the Southern mainland was very different from the repression strategy adopted in Sicily, mainly because of the different segments of the population being targeted by the military measures in the two contexts (only the peasants and the lower class in the former case, about the entire local community in the latter).



1867 by 6 percentage points, a 7% reduction relative to a mean of 0.809. The relative decrease in voter turnout in 1867 in repressed towns, interestingly, does not coincide with a statistically significant change in the concentration of votes across candidates running in the same national election (Column 6). This result plausibly suggests a minimal influence of the Sicilian mafia on general elections in the first years of the Kingdom of Italy.<sup>56</sup>

In Table 4, Column 7, I assess the impact of the repression campaign on trust in the central government in a larger segment of Sicilian society, by using the ratio of vaccinated children over population as outcome variable. I focus on local attitudes towards smallpox vaccination, a key instrument for prevention of a disease claiming over 400,000 people every year in Europe during the 18th century (Kopel, 2022). I find that Sicilian families living in towns that directly experienced government repression in 1863 were less likely to have their children vaccinated against smallpox in the period 1880-1884.

The estimated coefficient in Table 4, Column 7, is indicative of higher distrust in government in repressed towns, for several reasons. First, by the 1880s the Sicilian population was well aware of the existence of the smallpox vaccine. Smallpox prevention in Sicily began in 1801 under the Bourbon monarchy, with the first vaccination campaign being performed in Palermo (Chircop, 2010).<sup>57</sup> Second, the vaccine was distributed free of charge from the beginning, it became increasingly safe over time,<sup>58</sup> and its supply was based on a widespread distribution system, reaching

---

<sup>56</sup>A significant involvement of the Sicilian mafia in election manipulation started only in the 1870s, when the factional struggle to gain control of some of the state's power in Sicily became more intense (Dickie, 2004, Chapter 2). The new role of the mafia as protector of Sicilian politicians was explicitly revealed by a former chief prosecutor at the Palermo Court of Appeal, Diego Tajani, during a parliamentary debate in 1875, when he stated that “the mafia in Sicily [...] is dangerous and invincible because it is an instrument of local government” (quoted in Dickie, 2004, p. 73).

<sup>57</sup>In that year, the Spanish ruler made a special effort to protect the local population against this deadly disease, by organising the first free large-scale mass immunisation programme ever conducted in Italy (Bifulco et al., 2022, p. 3452). However, vaccination was never made compulsory in the Kingdom of the Two Sicilies, presumably owing to the Vatican's influence (Chircop, 2010, p. 169).

<sup>58</sup>In 1885 the smallpox vaccine was 94.3% effective after one dose, according to data for the whole Kingdom of Italy (Direzione Generale della Statistica, 1886*a*, p. CLXIV).

Table 4: Mechanism Analysis

Dependent variable:	Turnout 1861	Turnout 1865	Turnout 1867	Vote concentr. HHI 1861	Vote concentr. HHI 1865	Vote concentr. HHI 1867	Vaccinated children 1880-1884 p.c.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Repression 1863	-0.031 (0.019)	-0.013 (0.029)	-0.060*** (0.017)	-0.051 (0.048)	-0.002 (0.046)	0.006 (0.055)	-0.017** (0.008)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State presence controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	350	350	350	346	350	350	349
No. of clusters	24	24	24	24	24	24	24
P-value of wild cluster bootstrap test	0.121	0.654	0.004	0.280	0.971	0.916	0.062
R <sup>2</sup>	0.110	0.181	0.327	0.131	0.160	0.178	0.120
Mean of dependent variable	0.838	0.795	0.809	0.693	0.534	0.634	0.141

*Notes:* OLS estimates of the impact of exposure to government repression in 1863 on different proxies for trust in the central government at the town level. The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. The dependent variable in Columns 1 to 3 is the electoral turnout in, respectively, the 1861, 1865 and 1867 parliamentary elections. The dependent variable in Columns 4 to 6 is the Herfindahl index of the concentration of votes in the same elections. The dependent variable in Column 7 is the total number of children vaccinated between 1880 and 1884 over population in 1881. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo, distance from department capital and distance from polling station (removed in Column 1). State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861 (removed in Columns 1, 4, 5 and 6), dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. The p-value of the wild cluster bootstrap test of the hypothesis  $Repr1863 = 0$  is reported (bootstrap replications = 999). Standard errors clustered by department are reported in parentheses. \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

every town on the island.<sup>59</sup> My findings are hence unlikely to be explained by changes in vaccine supply. Third, the rulers, especially during the Bourbon regime, implemented specific strategies of health education to discourage the anti-vaccine front as much as possible. Parish priests were also frequently invited by the ruling elites to raise awareness among the population about the usefulness of vaccination (Chircop, 2010; Bifulco et al., 2022). Fourth, and crucially, until 1888 school-age children were the only social group required by law to be vaccinated against smallpox if they were to be admitted to elementary school (Della Peruta, 1980).<sup>60</sup> The data on vaccination rates among children thus provide a unique opportunity to study the local community’s reaction to a government rule and associate variations in the vaccine take-up rate with variations in trust in government. This relationship is made possible by the specific features of the smallpox vaccination campaign mentioned above, which imply that the observed mistrust is specific to trust in institutions and does not extend to the medical domain.<sup>61</sup>

Finally, an alternative expression of distrust in institutions could be provided by migration outflows. The repression campaign and the consequent decline in state legitimacy might have led people in repressed towns to ‘vote with their feet’ and move away from their hometown in Sicily. In Table 5 I find empirical evidence against this possible scenario. The relationship between the exposure to government repression

---

<sup>59</sup>Chircop (2010, p. 163) describes the smallpox vaccine distribution system under Bourbon rule as follows: “In Sicily, the whole project was managed by the central commission for vaccination [...] in Palermo, which was established by the Royal Decree of the 20th of October 1818 and operated through an island-wide network of *commissioni* at the provincial and communal levels [...], with the capillary ends of the network being the local vaccination boards (*giunte di vaccinazione*). Each *giunta* was made up of the mayor, any other state functionary on the spot, as well as the vaccinator and the parish priest. Their main responsibility lay in vaccinating as many inhabitants as possible in their localities.”

<sup>60</sup>In 1888 the enactment of the Crispi-Pagliani Law made smallpox vaccination compulsory for every newborn throughout Italy.

<sup>61</sup>A natural question is whether the alleged decline in state legitimacy due to the repression campaign led also to higher rates of tax evasion in treated towns. However, town-level data on tax compliance are unfortunately not available for Sicily around the Italian unification period. On the effects of Italy’s unification on taxpayer compliance across the Kingdom of Italy, see Acconcia et al. (2022).

Table 5: Ruled-out Explanations

Dependent variable:	Emigrants 1888-1892 p.c.	Ln pop. density 1871	Ln pop. density 1881
	(1)	(2)	(3)
Repression 1863	-0.003*** (0.001)	-0.013 (0.029)	-0.013 (0.034)
Province FE	Yes	Yes	Yes
Main controls	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes
State presence controls	Yes	Yes	Yes
N	349	350	350
No. of clusters	24	24	24
P-value of wild cluster bootstrap test	0.006	0.679	0.706
R <sup>2</sup>	0.304	0.987	0.985
Mean of dependent variable	0.002	4.900	5.017

*Notes:* OLS estimates of the impact of exposure to government repression in 1863 on migration outflows (Column 1) and the log of population density in 1871 (Column 2) and in 1881 (Column 3), all defined at the town level. The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. The dependent variable in Column 1 is the average number of permanent emigrants per capita in the period 1888-1892. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo, distance from department capital and distance from polling station (removed in Column 1). State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861 (removed in Columns 1, 4, 5 and 6), dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. The p-value of the wild cluster bootstrap test of the hypothesis  $Repr1863 = 0$  is reported (bootstrap replications = 999). Standard errors clustered by department are reported in parentheses. \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

in 1863 and the per capita amount of permanent emigrants leaving Sicily in the period 1888-1892 is negative and strongly significant (Column 1), meaning that the people living in repressed communities were actually *less* likely to move abroad in the medium term. This result is unlikely to be related to changes in local economic activity that occurred after 1863, as the effect of repression on town-level population density in 1871 and 1881 is found to be very small and insignificant (Columns 2 and 3). I interpret these results as suggesting that the population living in repressed towns eventually found a way to live with the oppressive public authorities, without necessarily leaving its hometown. The Sicilian mafia, able to enforce contracts and resolve disputes on demand, likely represented for these people the best local

alternative to the authoritarian and culturally distant central government.<sup>62</sup>

## 5.2 The Supply of Mafia

The results shown in the previous section stress the importance of declining state legitimacy in shaping the early diffusion of the Sicilian mafia. In this section I investigate another possible mechanism underlying my main findings: an increase in the ‘supply’ of mafia, due to a larger number of draft evaders joining pre-existing criminal groups in areas repressed by the Italian government in 1863. More precisely, I test whether the 1863 repression campaign played a reinforcing role in the expansion of banditry that started soon after the introduction of conscription in Sicily in 1861. The underlying conjecture is that the Sicilian mafia may have emerged also because of a stronger ‘military arm’ available to use, increasingly made of young men running away from law enforcement officials to avoid enlisting.

Before assessing the impact of government repression on draft evasion, it is helpful to understand how the Italian army’s recruitment system functioned in the first years of the Kingdom of Italy. The draft policy that was implemented in Sicily in 1861 was based on the 1854 system of male conscription designed by General Alfonso La Marmora for the Kingdom of Sardinia-Piedmont, which was extended to the whole country upon the Italian unification. According to this draft law, the central government was expected to announce every year the number of draftees required at the national level and afterwards assign each department a quota of draftees proportional to its population. Department councils were later required to divide their conscription share among the districts (*mandamenti*) in their own department, with the same operation being performed in turn by district councils with the draftees to be allocated among their towns.

The conscription policy also mandated that young men were to be called to

---

<sup>62</sup>It is always possible that Sicilians being initially settled in repressed towns decided to relocate to another place within the Italian Kingdom. Although this seems unlikely, internal migration data at the town level for the period of interest are unfortunately not available to address this question.

military service the year they turned 20. After conscription, these men were drafted into a two-tier system through a lottery. Those with the lowest numbers joined the army's first category and had to serve actively for five years. The men with the highest numbers were instead enlisted in the second category and received only a short basic training arranged periodically, joining active military service only in war time (Rovinello, 2013, p. 482). The last stage of the recruitment process consisted of medical examinations, set up by each department council. On specific dates (normally during the winter season), conscripts of each department assembled in the department capital to perform a health check or provide written evidence in support of an exemption request.<sup>63</sup> Those who failed to appear before their respective draft council during the medical examinations were declared draft evaders (*renitenti*).<sup>64</sup>

To quantify the size of draft evasion in Sicily in the wake of Italian unification, I use newly digitised department-level data from a series of volumes, known as “Torre reports,” published by the Italian Ministry of War and released annually since 1864.<sup>65</sup> These reports include statistical tables with aggregate data on all the Sicilian men drafted for the army, starting from the cohort born in 1840, the first one to be called up in 1861.<sup>66</sup> I rely on this source to construct the draft evasion rate (DER) for each Sicilian department and cohort, defined as the total number of draft evaders divided

---

<sup>63</sup>See Recruitment Law No. 1676, March 20, 1854, articles 54 and 58. There existed a long list of conditions that granted exemption from military service to whoever wanted to legally avoid the draft. For instance, clerics and men with family members dependent on them for support were excused from military service. The richest conscripts were also allowed by the draft law to buy themselves out of the obligation, by hiring a substitute or paying a fee directly to the Ministry of War (Rovinello, 2013, p. 483).

<sup>64</sup>Conscripts used to be informed of the date and time of the medical examinations through printed placards posted on the walls of their town center several months earlier (Recruitment Law No. 1676, March 20, 1854, article 32), so it is unlikely that draft evasion episodes stemmed from the conscripts' unawareness of their duty to report for military service.

<sup>65</sup>A'Hearn et al. (2009) named these documents after General Federico Torre, the head of the Italian Directorate of Conscription and Manpower from 1861 to 1891. This source has been largely used to conduct anthropometric research on historical adult heights: see, e.g., Costanzo (1948), Federico (2003), Arcaleni (2006) and the same A'Hearn et al. (2009).

<sup>66</sup>The 1840 and 1841 cohorts were summoned almost simultaneously at the end of 1861 with two separate military draft calls. In the following years the draft calls were organised annually and called one cohort at a time, so that the 1862 draft call was made on the cohort born in 1842, the 1863 draft call was made on the cohort born in 1843, and so on.

by the total number of men eligible for military service in each cohort, from 1861 to 1875.<sup>67</sup> Figure 1 shows the draft evasion rate across departments in Sicily for the first three cohorts that were ordered to enlist. In Western Sicily, and predominantly in the departments of Trapani and Palermo, resistance to the new draft policy started soon after its proclamation and persisted until the beginning of 1863, with evasion rates reaching values up to 40% or more after the third draft call.

To assess the effect of government repression across departments, I introduce a continuous measure of exposure to government repression, i.e. the share of repressed people in each department in 1863, defined as follows:

$$ReprShare1863_d = \frac{\text{Population in repressed towns in department } d}{\text{Total population in department } d},$$

where the population data come from the 1861 census.<sup>68</sup> This variable, depicted in Figure 6, varies considerably across departments, from a minimum of 0.06 recorded in Central Sicily to a maximum of 0.92 in the area close to Mazara del Vallo (in the southwestern part of the island), where 5 out of 6 towns fell victim to repression. Overall, the geographic variation of *ReprShare1863* is broadly in line with the main goal of Govone’s military campaign, which was the search for draft evaders.

I estimate the effect of exposure to government repression in 1863 on the draft evasion rate across Sicilian departments through the following event study:

$$DER_{dt} = \theta_t ReprShare1863_d + \pi_d + \rho_t + \sigma_{dt}, \quad (4)$$

where  $DER_{dt}$  denotes the draft evasion rate in department  $d$  for the eligible cohort

---

<sup>67</sup>I keep those who were exempted from military service in the denominator of the draft evasion rate, due to data limitations. However, it is worth pointing out that many conscripts tried to escape enlistment also by simulating disease or self-mutilation (Duggan, 2007, p. 285), thus causing the official number of draft evaders reported in the Torre reports to be biased downward.

<sup>68</sup>The precise number of people directly affected by General Govone’s repressive measures in each Sicilian town is rarely included in the archival sources, so it is not feasible to systematically calculate the repressed share at the town level. On the other hand, the frequent use of military cordons to encircle entire towns at night potentially affected all individuals in town, thus supporting the methodology in the construction of this variable.

in year  $t$  (born in year  $t - 20$ ) between 1860 and 1875,  $ReprShare1863_d$  represents the share of citizens in department  $d$  who suffered from government repression, and  $\pi_d$  and  $\rho_t$  indicate department and year fixed effects, respectively. Standard errors are clustered by department.<sup>69</sup> The set of coefficients  $\theta_t$  includes the coefficients on the interactions between  $ReprShare1863_d$  and the year fixed effects. The omitted category is 1862, the last pre-repression year. With this approach I compare the change in DER in departments that were exposed to different levels of repression intensity before and after the repression campaign. Moreover, department and year fixed effects allow me to control for, in order, any time-invariant department-specific characteristic and any unobserved variable in specific years that affects the whole region equally.<sup>70</sup>

In Figure 7 I show the draft evasion rate in the period of interest for the departments that were affected by the 1863 military campaign (for which the variable  $ReprShare1863$  is positive) and for those that were not (for which the variable  $ReprShare1863$  equals 0). Two patterns emerge clearly. On the one hand, draft evasion rates went down throughout the region right after the repression campaign and almost never returned to their pre-repression levels. This was likely due to the substantial errors characterising the early conscription lists, which included females and dead people among the first alleged draft evaders. These inaccuracies were fixed during General Govone’s expedition in Central and Western Sicily first and General Medici’s military campaign in Eastern Sicily soon after. On the other hand, after 1863 draft evasion rates were on average lower in repressed departments than in non-repressed ones, and consistently so over time. This trend is verified by the event study results shown in Figure 8 and Table A11. I find a decline in DER (statistically

---

<sup>69</sup>The choice to cluster standard errors at the department level and the small number of existing departments (24) do not allow me to add other control variables, as I do in previous parts of the paper.

<sup>70</sup>Department fixed effects are particularly relevant in this empirical setting, since they can capture heterogeneity across the different draft councils that coordinated the draft call-ups in their respective department.



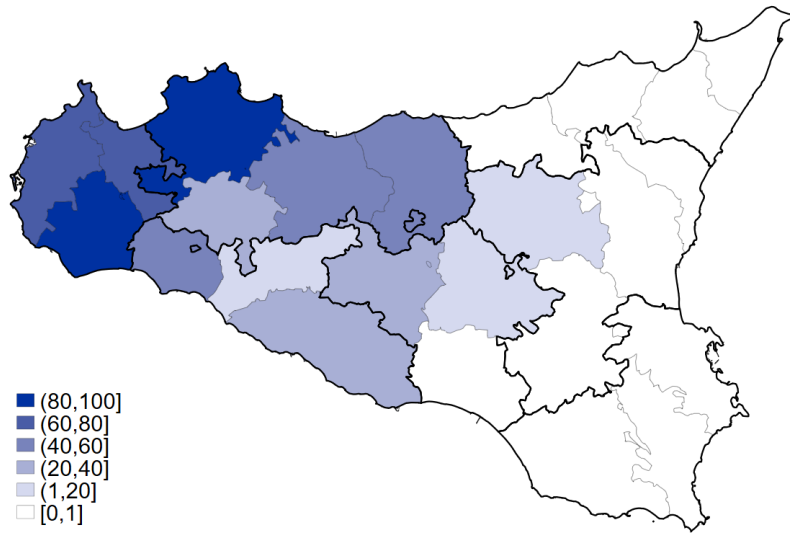
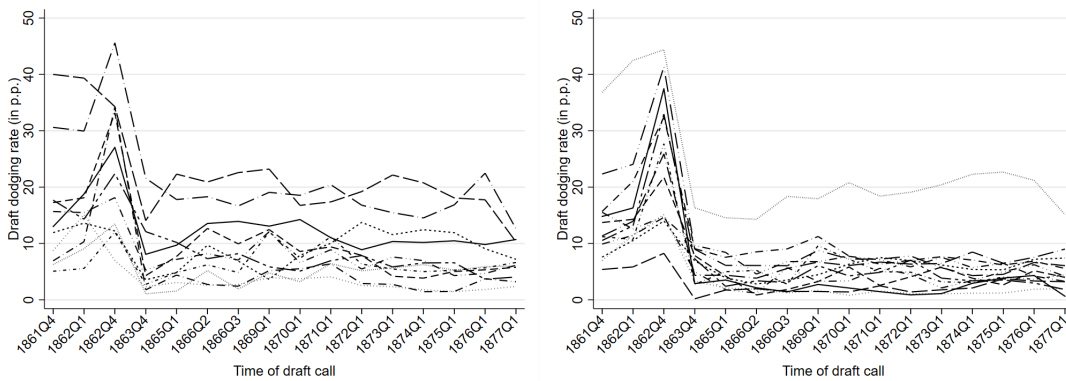


Figure 6: Share of citizens (in p.p.) repressed by the military in 1863 in each department (*circondario*). The methodology used to compute this variable is described in Section 5.2. The black lines represent province boundaries. *Sources*: 1861 census, Govone (1863), *Il Precursore* (1863), *La Forbice* (1863), *Il Giornale di Sicilia* (1863), Brancato (1956), Davis (1988) and Riall (1998).



(a) DER in non-repressed departments (for which  $ReprShare1863 = 0$ ).

(b) DER in repressed departments (for which  $ReprShare1863 > 0$ ).

Figure 7: Draft evasion rate (DER, in p.p.) at the department (*circondario*) level, by repression status. Each line represents a given department. *Sources*: Ministero della Guerra (1865-1877).



Figure 8: Estimated coefficients  $\theta_t$  from the event study model in Equation (4). Solid vertical lines plot the 95% confidence intervals, while the dashed vertical line indicates the time of the repression campaign.

significant at the 10% level) in departments more exposed to government repression in 1865, two years after the repression campaign, with the effect persisting until 1870. Non-significant estimates of the coefficients  $\theta_1$  and  $\theta_2$  lend support to the parallel trends assumption. In terms of magnitude, a department with an average share of repressed citizens (0.239) experienced a reduction of 1.2 percentage points in DER for the cohort born in 1845 (summoned in the second quarter of 1866), compared to a mean DER of 0.07 for the same cohort. I interpret these results as suggesting that the increase in brigandage following the introduction of conscription in Sicily may have not played a major role in the early spread of the Sicilian mafia, since government repression in 1863 had a deterrent effect on draft evasion ever since.<sup>71</sup>

Despite the empirical evidence in Figure 8, one should not completely rule out the supply mechanism, since the group of draft evaders reported on the conscription

<sup>71</sup>This is consistent with the hypothesis that people in repressed towns might have formed expectations on the central government's ability to resort to repression again in the future if facing new acts of local opposition, as theorised by Rozenas and Zhukov (2019).

lists represented only a fraction of the outlaws in Sicily. This is confirmed by Lupo (1996, p. 32), when he writes that “a huge number of individuals became fugitives from justice [...] as a direct result of the army’s actions.”<sup>72</sup> It is possible that these fugitives did not enter official criminal records, but still contributed to sustain the demand for private protection that fuelled the emergence of the mafia.

## 6 Persistence

In this section I replicate the exercise in Table 2 to investigate the persistence of the 1863 repression campaign into 1900. To do so I employ a town-level ordinal variable, *Mafia*1900, ranging from 0 (no mafia presence) to 3 (high mafia presence), using data from Cutrera (1900). I regress this new measure of historical mafia activity on the treatment variable *Repr*1863 and the same set of confounders included in the vector  $x_{ip}$  in Equations (1) and (2), with the addition of a dummy variable accounting for the presence of the Peasant Fasci socialist movement between 1891 and 1894, in the spirit of Acemoglu et al. (2020).

Two main reasons motivate this persistence analysis. First, given the extensive use of Cutrera’s data in previous studies on the early evolution of the Sicilian mafia (Gambetta, 1993; Buonanno et al., 2015; Dimico et al., 2017; Acemoglu et al., 2020), using this measure of mafia presence allows for a more direct comparison between my work and the existing literature on the mafia’s emergence. Second, by focusing on mafia presence in 1900 I can analyse the long-run effects of violent exposure to a coercive nation-building policy in Sicily across the second half of the 19th century.

Table 6 shows OLS and 2SLS estimates of the long-run relationship between exposure to government repression and mafia presence, always on the restricted sample. In Panel D the OLS estimate of  $\beta_1$  is positive throughout all regressions, but it loses statistical significance with the inclusion of other determinants of mafia

---

<sup>72</sup>Along the same lines Lupo (1996, p. 32) also adds the following statement taken from an anonymous Palermitan in 1867: “I have been reliably informed that in one village in this district, there were thirty-four fugitives from justice in proportion to a single draft evader.”

Table 6: Persistence Analysis

	(1)	(2)	(3)	(4)	(5)
Panel A: Second stage					
	Dependent variable: Mafia 1900				
Regression 1863	4.078*** (0.912)	1.834 (1.104)	1.765* (0.996)	1.370* (0.682)	1.221* (0.660)
Mean of dependent variable	1.404	1.404	1.404	1.404	1.404
Panel B: First stage					
	Dependent variable: Regression 1863				
ln(1 + Distance from least-cost path)	-0.118*** (0.018)	-0.114*** (0.028)	-0.111*** (0.029)	-0.123*** (0.030)	-0.119*** (0.030)
Kleibergen-Paap Wald F statistic	43.90	16.79	14.34	16.32	15.49
Panel C: Reduced form					
	Dependent variable: Mafia 1900				
ln(1 + Distance from least-cost path)	-0.482*** (0.060)	-0.209** (0.100)	-0.196** (0.084)	-0.168** (0.065)	-0.146** (0.066)
P-value of wild cluster bootstrap test	0.000	0.075	0.087	0.032	0.050
R <sup>2</sup>	0.333	0.439	0.512	0.562	0.592
Panel D: OLS Estimates					
	Dependent variable: Mafia 1900				
Regression 1863	1.132*** (0.194)	0.369*** (0.129)	0.240* (0.128)	0.198 (0.130)	0.174 (0.125)
P-value of wild cluster bootstrap test	0.001	0.058	0.118	0.152	0.206
R <sup>2</sup>	0.112	0.427	0.498	0.555	0.587
Province FE	No	Yes	Yes	Yes	Yes
Main controls	No	No	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes
State presence controls	No	No	No	No	Yes
N	277	277	277	277	277
No. of clusters	24	24	24	24	24

*Notes:* 2SLS (Panel A) and OLS (Panel D) estimates of the impact of exposure to government repression in 1863 on the intensity of mafia presence in 1900 at the town level. The dependent variable is an ordinal variable indicating the level of mafia activity in 1900 on a 0 to 3 scale (0 meaning no mafia presence, 3 meaning high mafia presence), according to Cutrera (1900). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. This regressor is instrumented by the natural logarithm of (one plus) the distance to the network of least-cost paths described in Section 4.3. The ten garrison towns used to construct this optimal route network are dropped from the sample. Panel B reports the results of the first-stage regression for the exposure to government repression, whereas Panel C includes estimates of the reduced-form regression of mafia presence on the excluded instrument. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870, dummy for banditry presence in 1820-1849 and dummy for Peasant Fasci presence in 1891-1894. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. Standard errors clustered by department are reported in parentheses. The p-value of the wild cluster bootstrap test of the hypothesis  $Repr_{1863} = 0$  is reported (bootstrap replications = 999). \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

presence in the estimating equation. The IV second-stage results in Panel A are instead both larger in magnitude and statistically significant at the 10% level even in the presence of province fixed effects. The estimate from Panel A, Column 5, indicates that being exposed to government repression in 1863 increases the mafia index in 1900 in a given treated town by approximately 1.2. This result resembles the estimated impact of the Peasant Fasci on the strength of the Sicilian mafia found by Acemoglu et al. (2020) in their IV estimations. At the same time, the effect I estimate by 2SLS is almost 2.5 times larger than the one identified by Buonanno et al. (2015) and Dimico et al. (2017) when they use similar estimation strategies.<sup>73</sup> Overall, these findings suggest that the government repression episodes of 1863 left notable marks on Sicilian history and represented a key driver of the spread of the Sicilian mafia both in the medium and long term.<sup>74</sup>

## 7 Conclusion

Can coercive nation building lead to the emergence of mafia-type organisations? In this paper I empirically investigate this question by studying the origins of the Sicilian mafia. Many studies have proposed theories of the early spread of this criminal syndicate (Bandiera, 2003; Buonanno et al., 2015; Dimico et al., 2017; Acemoglu et al., 2020), in which the institutional context is often exogenously determined and the mafia's power is either a function of permanent structural factors (e.g. citrus suitability, sulphur presence, land fragmentation), or it results from historical accidents (as the exceptional drought in 1893). This paper complements the existing literature on the emergence of the Sicilian mafia by exploring the direct role of

---

<sup>73</sup>Comparison based on estimates reported in Buonanno et al. (2015), Table 4, and Dimico et al. (2017), Table 5, Column 1.

<sup>74</sup>Since the proxy for mafia presence used in this analysis is an ordinal variable, I also estimate an ordered logit model with department fixed effects, using the so-called 'blow-up and cluster' estimator introduced by Baetschmann et al. (2015). The corresponding Stata command `feologit` is discussed in Baetschmann et al. (2020). Table A10 reports the results of this estimation, which are broadly in line with the main findings shown in Table 6.

the central Italian government when it unintentionally fostered the mafia's initial development.

I show that the early rise of organised crime in Sicily can be partly credited to the central government's failed efforts to instil a sense of national identity among the Sicilian population. In particular, I focus on the instances of cultural resistance to forced assimilation that followed the introduction of conscription in Sicily in 1861. The conscription policy, aiming at "making the Italians," generated widespread demand for private protection in Sicilian society, especially after a repressive military campaign run by the new central government in 1863 to tackle high rates of draft evasion on the island. The repression campaign boosted the demand for protection from law enforcers, which was readily met by the mafia, the only trusted provider of protection in Sicily after 1863.

Using a variety of identification strategies, including an IV approach, I provide suggestive evidence that Sicilian towns victimised by the repression campaign were more likely to experience the presence of organised crime groups soon after the military expedition and also many decades later. I interpret this result as evidence of decreasing state legitimacy following the forced implementation of the conscription policy.

Although this work is directly informative only about 19th-century Sicily, its findings could apply to other contexts as well, given the widespread introduction of coercive nation-building policies especially in recent years (Human Rights Watch, 2021; Said-Moorhouse and Ochman, 2022; Dehdari and Gehring, 2022). Overall, this paper supports the main conclusion reached by Acemoglu and Jackson (2017). They claim that policies in conflict with prevailing social norms can backfire, inducing more people to break the law. Drawing a parallel with this study, my results suggest that nation-building policies that are not mindful of the cultural context of the targeted social group can encourage law-breaking in society and fuel organised crime.

## References

- Abadie, A. and Imbens, G. W. (2006), ‘Large Sample Properties of Matching Estimators for Average Treatment Effects’, *Econometrica*, **74**(1), 235–267.
- Abadie, A. and Imbens, G. W. (2011), ‘Bias-Corrected Matching Estimators for Average Treatment Effects’, *Journal of Business and Economic Statistics*, **29**(1), 1–11.
- Accetturo, A., Bugamelli, M. and Lamorgese, A. (2017), ‘Law Enforcement and Political Participation: Italy, 1861–65’, *Journal of Economic Behavior and Organization*, **140**(C), 224–245.
- Acconcia, A., D’Amato, M., Martina, R. and Ratto, M. (2022), ‘The Response of Taxpayer Compliance to the Large Shock of Italian Unification’, *European Journal of Political Economy*, **73**, 102158.
- Acemoglu, D., Cantoni, D., Johnson, S. and Robinson, J. A. (2011), ‘The Consequences of Radical Reform: The French Revolution’, *American Economic Review*, **101**(7), 3286–3307.
- Acemoglu, D., Feo, G. D. and Luca, G. D. D. (2020), ‘Weak States: Causes and Consequences of the Sicilian Mafia’, *Review of Economic Studies*, **87**(2), 537–581.
- Acemoglu, D. and Jackson, M. O. (2017), ‘Social Norms and the Enforcement of Laws’, *Journal of the European Economic Association*, **15**(2), 245–295.
- Acemoglu, D. and Robinson, J. A. (2021), Culture, Institutions and Social Equilibria: A Framework. NBER Working Paper No. 28832.
- Aghion, P., Jaravel, X., Persson, T. and Rouzet, D. (2019), ‘Education and Military Rivalry’, *Journal of the European Economic Association*, **17**(2), 376–412.
- A’Hearn, B., Peracchi, F. and Vecchi, G. (2009), ‘Height and the Normal Distribution: Evidence from Italian Military Data’, *Demography*, **46**(1), 1–25.
- Alatri, P. (1954), *Lotte Politiche in Sicilia Sotto il Governo della Destra (1866-1874)*, Turin, IT: Einaudi.
- Alesina, A., Baqir, R. and Easterly, W. (1999), ‘Public Goods and Ethnic Divisions’, *Quarterly Journal of Economics*, **114**(4), 1243–1284.
- Alesina, A. and Giuliano, P. (2015), ‘Culture and Institutions’, *Journal of Economic Literature*, **53**(4), 898–944.
- Alesina, A., Giuliano, P. and Reich, B. (2021), ‘Nation-Building and Education’, *Economic Journal*, **131**(638), 2273–2303.

- Alesina, A., Reich, B. and Riboni, A. (2020), ‘Nation-Building, Nationalism, and Wars’, *Journal of Economic Growth*, **25**(4), 381–430.
- Alongi, G. (1977), *La Maffia*, Palermo, IT: Sellerio.
- Altonji, J., Elder, T. and Taber, C. R. (2005), ‘An Evaluation of Instrumental Variable Strategies for Estimating the Effects of Catholic Schooling’, *Journal of Human Resources*, **40**(4), 791–821.
- Amodio, F. (2012), ‘Hard to Forget: Long-Lasting Effects of Social Capital Accumulation Shocks’. AICCON Working Paper No. 105.
- Arcaleni, E. (2006), ‘Secular Trend and Regional Differences in the Stature of Italians, 1854-1980’, *Economics and Human Biology*, **4**(1), 24–38.
- Arlacchi, P. (1983), *Mafia, Peasants, and Great Estates*, Cambridge, UK: Cambridge University Press.
- Assouad, L. (2020), ‘Charismatic Leaders and Nation Building’. PSE Working Papers halshs-02873520, HAL.
- Baetschmann, G., Ballantyne, A., Staub, K. E. and Winkelmann, R. (2020), ‘feologit: A New Command for Fitting Fixed-Effects Ordered Logit Models’, *Stata Journal*, **20**(2), 253–275.
- Baetschmann, G., Staub, K. E. and Winkelmann, R. (2015), ‘Consistent Estimation of the Fixed-Effects Ordered Logit Model’, *Journal of the Royal Statistical Society Series A*, **178**(3), 685–703.
- Bandiera, O. (2003), ‘Land Reform, the Market for Protection, and the Origins of the Sicilian Mafia: Theory and Evidence’, *Journal of Law, Economics, and Organization*, **19**(1), 218–244.
- Bandiera, O., Mohnen, M., Rasul, I. and Viarengo, M. (2019), ‘Nation-Building Through Compulsory Schooling during the Age of Mass Migration’, *Economic Journal*, **129**(617), 62–109.
- Bargain, O. and Aminjonov, U. (2020), ‘Trust and Compliance to Public Health Policies in Times of COVID-19’, *Journal of Public Economics*, **192**, 104316.
- Basile, R., Ciccarelli, C. and Groote, P. (2022), ‘The Legacy of Literacy: Evidence from Italian Regions’, *Regional Studies*, **56**(5), 794–807.
- Bazzi, S., Gaduh, A., Rothenberg, A. D. and Wong, M. (2019), ‘Unity in Diversity? How Intergroup Contact Can Foster Nation Building’, *American Economic Review*, **109**(11), 3978–4025.



- Bifulco, M., Zazzo, E. D., Pisanti, S., Martini, M. and Orsini, D. (2022), ‘The Nineteenth-Century Experience of the Kingdom of the Two Sicilies on Mandatory Vaccination: An Italian Phenomenon?’, *Vaccine*, **40**(25), 3452–3454.
- Bingley, P. and Martinello, A. (2017), ‘Measurement Error in Income and Schooling and the Bias of Linear Estimators’, *Journal of Labor Economics*, **35**(4), 1117–1148.
- Bisin, A., Patacchini, E., Verdier, T. and Zenou, Y. (2011), ‘Formation and Persistence of Oppositional Identities’, *European Economic Review*, **55**(8), 1046–1071.
- Bisin, A. and Verdier, T. (2017), On the Joint Evolution of Culture and Institutions. NBER Working Paper No. 23375.
- Blair, R. A., Morse, B. S. and Tsai, L. L. (2017), ‘Public Health and Public Trust: Survey Evidence from the Ebola Virus Disease Epidemic in Liberia’, *Social Science & Medicine*, **172**, 89–97.
- Blok, A. (1975), *The Mafia of a Sicilian Village, 1860-1960: A Study of Violent Peasant Entrepreneurs (State and Revolution)*, New York: Harper & Row.
- Blouin, A. and Mukand, S. W. (2019), ‘Erasing Ethnicity? Propaganda, Nation Building, and Identity in Rwanda’, *Journal of Political Economy*, **127**(3), 1008–1062.
- Bogart, D., You, X., Alvarez-Palau, E. J., Satchell, M. and Shaw-Taylor, L. (2022), ‘Railways, Divergence, and Structural Change in 19th Century England and Wales’, *Journal of Urban Economics*, **128**, 103390.
- Brancato, F. (1956), *Storia della Sicilia Post-Unificazione. Parte Prima: La Sicilia nel Primo Ventennio del Regno d’Italia*, Bologna, IT: Cesare Zuffi Editore.
- Buonanno, P., Durante, R., Prarolo, G. and Vanin, P. (2015), ‘Poor Institutions, Rich Mines: Resource Curse in the Origins of the Sicilian Mafia’, *Economic Journal*, **125**(586), F175–F202.
- Cancila, O. (1984), *Come Andavano le Cose nel Sedicesimo Secolo*, Palermo, IT: Sellerio.
- Cannella, M., Makarin, A. and Pique Cebreco, R. S. (2021), ‘The Political Legacy of Nazi Annexation’. Available at SSRN: <https://ssrn.com/abstract=3817933> or <http://dx.doi.org/10.2139/ssrn.3817933>.
- Carbone, S. and Grispo, R. (1969), *L’Inchiesta sulle Condizioni Sociali ed Economiche della Sicilia (1875-1876), Volume 2*, Bologna, IT: Cappelli.

- Carraro, L. (1976), *Relazione Conclusiva della Commissione Parlamentare d'Inchiesta sul Fenomeno della Mafia in Sicilia*, Rome, IT: Tipografia del Senato.
- Carvalho, J.-P. (2013), 'Veiling', *Quarterly Journal of Economics*, **128**(1), 337–370.
- Carvalho, J.-P. and Koyama, M. (2016), 'Resisting Education'. GMU Working Paper in Economics No. 13-18.
- Cary, J. (1799), 'A New Map of Italy, Including the Islands of Sicily, Sardinia and Corsica with the Post Roads from the Latest Authorities'. London, UK: Cary Engraver and Map-seller.
- Chircop, J. (2010), "Giusta la Benefica Intenzione del Re": The Bourbon Cowpox Vaccination Campaign in Sicily', *Hygiea Internationalis*, **9**(1), 155–181.
- Ciccarelli, C., Dalmazzo, A. and Razzolini, T. (2021), 'Sicilian Sulphur and Mafia: Resources, Working Conditions and the Practice of Violence'. IZA Discussion Paper No. 14959.
- Clark, M. (2009), *The Italian Risorgimento*, Routledge.
- Clots-Figueras, I. and Masella, P. (2013), 'Education, Language and Identity', *Economic Journal*, **123**(570), F332–F357.
- Colajanni, N. (1895), *Gli Avvenimenti di Sicilia e le Loro Cause*, Palermo, IT: Remo Sandron.
- Colajanni, N. (1900), *Nel Regno della Mafia. Dai Borboni ai Sabaudi*, Rome, IT: Rivista Popolare.
- Colella, F., Lalive, R., Sakalli, S. O. and Thoenig, M. (2019), 'Inference with Arbitrary Clustering'. IZA Discussion Paper No. 12584.
- Conley, T. (1999), 'GMM Estimation with Cross Sectional Dependence', *Journal of Econometrics*, **92**(1), 1–45.
- Costanzo, A. (1948), 'La Statura degli Italiani Ventenni Nati dal 1854 al 1920', *Annali di Statistica, Serie VIII*, **2**, 63–123.
- Cutrera, A. (1900), *La Mafia e i Mafiosi*, Palermo, IT: Reber.
- Damiani, A. (1885), *Relazione del Commissario Abele Damiani, Deputato al Parlamento, sulla Prima Circostrizione (Province di Caltanissetta, Catania, Girgenti, Messina, Palermo, Siracusa e Trapani)*, in S. Jacini, ed., 'Atti della Giunta per la Inchiesta Agraria e sulle Condizioni della Classe Agricola', Rome, IT: Forzani e C.

- Davis, J. A. (1988), *Conflict and Control: Law and Order in Nineteenth-Century Italy*, London: Palgrave Macmillan.
- De Jong, M. and Mamadouh, V. (2002), *Two Contrasting Perspectives on Institutional Transplantation*, in M. De Jong, K. Lalenis and V. Mamadouh, eds, 'The Theory and Practice of Institutional Transplantation: Experiences with the Transfer of Policy Institutions', Springer Dordrecht.
- De Sanctis, G. (1840), *Dizionario Statistico de' Paesi del Regno delle Due Sicilie*, Naples, IT: Stamperia e cartiere del Fibreno.
- Dehdari, S. H. and Gehring, K. (2022), 'The Origins of Common Identity: Evidence from Alsace-Lorraine', *American Economic Journal: Applied Economics*, **14**(1), 261–292.
- Del Monte, A. and Pennacchio, L. (2012), 'Agricultural Productivity, Banditry and Criminal Organisations in Post-Unification Italy', *Rivista Italiana degli Economisti*, **17**(3), 347–377.
- Della Peruta, F. (1980), 'Sanità Pubblica e Legislazione Sanitaria dall'Unità a Crispi', *Studi Storici*, **21**(4), 713–759.
- Depetris-Chauvin, E., Durante, R., and Campante, F. (2020), 'Building Nations through Shared Experiences: Evidence from African Football', *American Economic Review*, **110**(5), 1572–1602.
- Desmet, K., Ortuño-Ortín, I. and Wacziarg, R. (2017), 'Culture, Ethnicity, and Diversity', *American Economic Review*, **107**(9), 2479–2513.
- Dickie, J. (2004), *Cosa Nostra: A History of the Sicilian Mafia*, New York: Palgrave Macmillan.
- Dimico, A., Isopi, A. and Olsson, O. (2017), 'Origins of the Sicilian Mafia: The Market for Lemons', *Journal of Economic History*, **77**(4), 1083–1115.
- Direzione Generale della Statistica (1865), *Statistica del Regno d'Italia. Istruzione Elementare Pubblica per Comuni. Anno Scolastico 1862-1863*, Modena, IT: Antonio ed Angelo Cappelli.
- Direzione Generale della Statistica (1867), *Statistica del Regno d'Italia. Elezioni Politiche e Amministrative. Anni 1865-1866*, Florence, IT: Tofani.
- Direzione Generale della Statistica (1886a), *Risultati dell'Inchiesta sulle Condizioni Igieniche e Sanitarie nei Comuni del Regno. Parte Prima*, Rome, IT: Ospizio di S. Michele.

- Direzione Generale della Statistica (1886b), *Risultati dell'Inchiesta sulle Condizioni Igieniche e Sanitarie nei Comuni del Regno. Parte Seconda*, Rome, IT: Ospizio di S. Michele.
- Dryzek, J. S. (1996), *The Informal Logic of Institutional Design*, in R. E. Goodin, ed., 'The Theory of Institutional Design', Cambridge, UK: Cambridge University Press.
- Duggan, C. (2007), *The Force of Destiny: A History of Italy Since 1796*, London: Penguin Books.
- Durante, R., Guiso, L. and Gulino, G. (2021), 'Asocial Capital: Civic Culture and Social Distancing during COVID-19', *Journal of Public Economics*, **194**, 104342.
- Esposito, E., Rotesi, T., Saia, A. and Thoenig, M. (2021), 'Reconciliation Narratives: The Birth of a Nation After the US Civil War'. CEPR Press Discussion Paper No. 15938.
- Federico, G. (2003), 'Heights, Calories and Welfare: A New Perspective on Italian Industrialization, 1854-1913', *Economics and Human Biology*, **1**(3), 289–308.
- Fenske, J., Kala, N. and Wei, J. (2021), 'Railways and Cities in India'. CAGE Working Paper No. 559.
- Fernández, R. (2011), *Chapter 11 - Does Culture Matter?*, in J. Benhabib, A. Bisin and M. O. Jackson, eds, 'Handbook of Social Economics', Vol. 1, Amsterdam: North-Holland, pp. 481–510.
- Fiume, G. (1984), *Le Bande Armate in Sicilia (1819-1849): Violenza e Organizzazione del Potere*, Palermo, IT: Annali della Facoltà di Lettere e Filosofia dell'Università di Palermo.
- Fouka, V. (2020), 'Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I', *Review of Economic Studies*, **87**(1), 204–239.
- Franchetti, L. (1877), *Condizioni Politiche e Amministrative della Sicilia*, Florence, IT: Barbera.
- Gambetta, D. (1993), *The Sicilian Mafia: The Business of Private Protection*, Cambridge, MA: Harvard University Press.
- Gorodnichenko, Y. and Roland, G. (2011), 'Which Dimensions of Culture Matter for Long-Run Growth?', *American Economic Review*, **101**(3), 492–498.
- Guiso, L., Sapienza, P. and Zingales, L. (2000), 'The Role of Social Capital in Financial Development'. NBER Working Paper No. 7563.

- Hess, H. (1973), *Mafia and Mafiosi: The Structure of Power*, Farnborough, UK: Saxon House.
- Hornung, E. (2015), 'Railroads and Growth in Prussia', *Journal of the European Economic Association*, **13**(4), 699–736.
- Human Rights Watch (2021), 'Break Their Lineage, Break Their Roots'. April 19 [online]. Available at: <https://www.hrw.org/report/2021/04/19/break-their-lineage-break-their-roots/chinas-crimes-against-humanity-targeting>.
- Jedwab, R. and Moradi, A. (2016), 'The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa', *Review of Economics and Statistics*, **98**(2), 268–284.
- Komisarchik, M., Sen, M. and Velez, Y. R. (2020), 'The Political Consequences of Ethnically Targeted Incarceration: Evidence from Japanese-American Internment During WWII'. HKS Working Paper RWP20-021.
- Kopel, J. (2022), 'An Exploration of Vaccination in the 19th Century Through the Eyes of Dr. Albert Mackey', *Baylor University Medical Center Proceedings*, **35**(4), 505–509.
- Kopsidis, M. and Bromley, D. W. (2016), 'The French Revolution and German Industrialization: Dubious Models and Doubtful Causality', *Journal of Institutional Economics*, **12**(1), 161–190.
- Larson, H. J. and Heymann, D. L. (2010), 'Public Health Response to Influenza A(H1N1) as an Opportunity to Build Public Trust', *Journal of the American Medical Association*, **303**(3), 271–272.
- Lecce, G. and Ogliari, L. (2019), 'Institutional Transplant and Cultural Proximity: Evidence from Nineteenth-Century Prussia', *Journal of Economic History*, **79**(4), 1060–1093.
- Lecce, G., Ogliari, L. and Orlando, T. (2022), 'State Formation, Social Unrest and Cultural Distance', *Journal of Economic Growth*, **27**(3), 453–483.
- Lo Jacono, V. (1856), *Indice Generale dei Comuni, Sottocomuni, Casali, Villaggi, Borgate ed Isole della Sicilia con la indicazione delle distanze dalla Capitale, dal Capoluogo della Provincia e del Distretto*, Palermo, IT: Tipografia del Commercio.
- Lorenzoni, G. (1910), *Inchiesta Parlamentare sulle Condizioni dei Contadini nelle Provincie Meridionali e nella Sicilia. Vol. VI - Sicilia, Tomo I*, Rome, IT: Tipografia Nazionale Bertero.

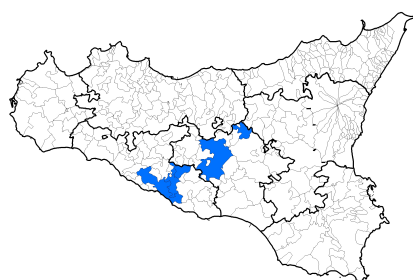
- Lowes, S. and Montero, E. (2021), ‘The Legacy of Colonial Medicine in Central Africa’, *American Economic Review*, **111**(4), 1284–1314.
- Lowes, S., Nunn, N., Robinson, J. A. and Weigel, J. L. (2017), ‘The Evolution of Culture and Institutions: Evidence From the Kuba Kingdom’, *Econometrica*, **85**(4), 1065–1091.
- Lupo, S. (1996), *Storia della Mafia*, Rome, IT: Donzelli Editore.
- Mack Smith, D. (1968), *A History of Sicily: Volume 2, Modern Sicily After 1713*, New York: Dorset Press.
- Martinez-Bravo, M. and Stegmann, A. (2022), ‘In Vaccines We Trust? The Effects of the CIA’s Vaccine Ruse on Immunization in Pakistan’, *Journal of the European Economic Association*, **20**(1), 150–186.
- Maurici, A. (1915), *Il Regime Dispotico del Governo d’Italia in Sicilia dopo Aspromonte, Settembre 1862 - Dicembre 1863*, Palermo, IT: E. Priulla.
- Miguel, E. and Gugerty, M. K. (2005), ‘Ethnic Diversity, Social Sanctions, and Public Goods in Kenya’, *Journal of Public Economics*, **89**(11-12), 2325–2368.
- Ministero della Guerra (1864), *Relazione al Sig. Ministro della Guerra sulle Leve Eseguite in Italia dalle Annessioni delle Varie Provincie al 30 Settembre 1863*, Turin, IT: Tipografia Fodratti.
- Ministero della Guerra (1865-1877), *Della Leva sui Giovani Nati nell’Anno ... e delle Vicende dell’Esercito dal 1° Ottobre ... al 30 Settembre .... Relazione del Maggiore Generale Federico Torre al Sig. Ministro della Guerra*, Turin, IT: Tipografia Fodratti.
- Ministero della Guerra (1872), *Regolamento di Disciplina Militare*, Rome, IT: C. Voghera.
- Oster, E. (2019), ‘Unobservable Selection and Coefficient Stability: Theory and Evidence’, *Journal of Business and Economic Statistics*, **37**(2), 187–204.
- Parodi, L. (1873), *Sull’Estrazione dello Zolfo in Sicilia e sugli Usi Industriali del Medesimo*, Florence, IT: Barbera.
- Pezzino, P. and Mangiameli, R. (2016), ‘Le Mafie tra Immaginazione e Realtà. Una Discussione a Proposito di F. Benigno, La Mala Setta. Intorno alle Origini di Mafia e Camorra, 1859-1878’, *Annali della Scuola Normale Superiore di Pisa, Classe di Lettere e Filosofia, serie 5*, **8**(1), 247–275.

- Redding, S. J. and Turner, M. A. (2015), *Transportation Costs and the Spatial Organization of Economic Activity*, in G. Duranton, J. V. Henderson and W. C. Strange, eds, 'Handbook of Regional and Urban Economics, Volume 5', New York: Elsevier.
- Renda, F. (1977), *I Fasci Siciliani (1892-1894)*, Turin, IT: Einaudi.
- Riall, L. (1998), *Sicily and the Unification of Italy. Liberal Policy and Local Power, 1859-1866*, Oxford, UK: Clarendon Press.
- Roland, G. (2004), 'Understanding Institutional Change: Fast-Moving and Slow-Moving Institutions', *Studies in Comparative International Development*, **38**(4), 109–131.
- Rovinello, M. (2013), *The Draft and Draftees in Italy, 1861-1914*, in E.-J. Zürcher, ed., 'Fighting for a Living: A Comparative Study of Military Labour 1500-2000', Amsterdam University Press.
- Rozenas, A., Schutte, S. and Zhukov, Y. M. (2017), 'The Political Legacy of Violence: The Long-Term Impact of Stalin's Repression in Ukraine', *Journal of Politics*, **79**(4), 1147–1161.
- Rozenas, A. and Zhukov, Y. M. (2019), 'Mass Repression and Political Loyalty: Evidence from Stalin's 'Terror by Hunger'', *American Political Science Review*, **113**(2), 569–583.
- Russo, N. (1964), *Antologia della Mafia*, Palermo, IT: Il Punto Edizioni.
- Said-Moorhouse, L. and Ochman, O. (2022), 'This is What the 'Russification' of Ukraine's Education System Looks Like in Occupied Areas'. CNN, May 16 [online]. Available at: <https://edition.cnn.com/2022/05/16/europe/russia-ukraine-education-intl-cmd/index.html>.
- Scott, J. C. (2009), *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*, New Haven, CT: Yale University Press.
- Shurkin, M., Gordon IV, J., Frederick, B. and Pernin, C. G. (2017), *Building Armies, Building Nations: Toward a New Approach to Security Force Assistance*, Santa Monica, CA: RAND Corporation.
- Smyth, G. E. (1826), 'Carta Generale della Isola di Sicilia Compilata Disegnata ed Incisa Nell'Ufficio Topografico di Napoli'. Bibliothèque Nationale de France, GED-4625. Available at: <https://gallica.bnf.fr/ark:/12148/btv1b53006629d> (last access: October 23, 2022).

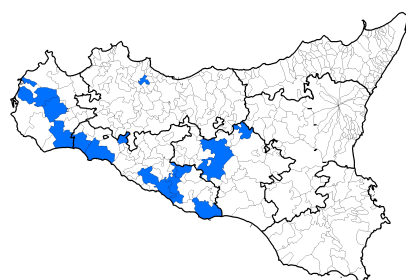
- Stato Maggiore del Regio Esercito (1863), 'Storia dei Corpi - Fanteria', Volume 7. Ufficio Storico dello Stato Maggiore dell'Esercito, Rome, IT.
- Sylos Labini, P. (2003), *Le Radici della Mafia in Sicilia*, in P. Sylos Labini and G. Arena, eds, 'Scritti sul Mezzogiorno (1954-2001)', Manduria, IT: Piero Lacaita.
- Słoczyński, T. (2022), 'Interpreting OLS Estimands when Treatment Effects are Heterogeneous: Smaller Groups Get Larger Weights', *Review of Economics and Statistics*, **104**(3), 501–509.
- Tur-Prats, A. and Valencia Caicedo, F. (2020), 'The Long Shadow of the Spanish Civil War'. CEPR Press Discussion Paper No. 15091.
- Ulloa, P. C. (1838), *Considerazioni sullo Stato Economico e Politico della Sicilia*, in E. Pontieri, ed., 'Il Riformismo Borbonico nella Sicilia del Sette e dell'Ottocento', 1965, Naples, IT: Edizioni Scientifiche Italiane.
- You, Y. (2018), 'Unified Language, Labor and Ideology'. Mimeo, Harvard University. Available at: [https://scholar.harvard.edu/files/yangyou/files/language\\_labor\\_ideology20180317.pdf](https://scholar.harvard.edu/files/yangyou/files/language_labor_ideology20180317.pdf).



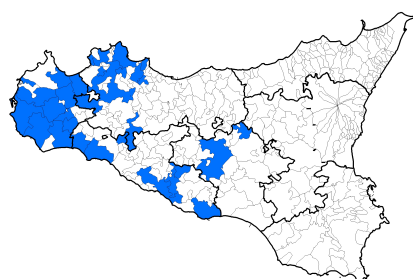
## Appendix



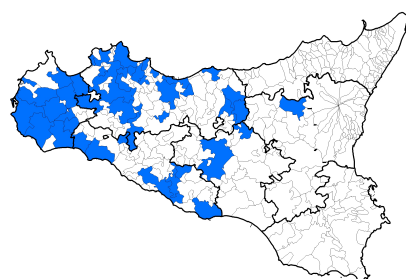
(a) June 26 - July 31.



(b) June 26 - August 31.



(c) June 26 - September 30.



(d) June 26 - November 11.

Figure A1: Evolution of the 1863 repression campaign over time. The blue polygons indicate towns that fell victim to government repression. The black lines represent province boundaries. *Sources:* Govone (1863), *Il Precursore* (1863), *La Forbice* (1863), *Il Giornale di Sicilia* (1863), Brancato (1956), Davis (1988), Riall (1998).

Table A1: Descriptive Statistics

Variable	N	Mean	SD	Min	Max
<b>Main variables</b>					
Mafia presence in 1875	350	0.123	0.329	0	1
Mafia presence in 1900	287	1.422	1.134	0	3
Government repression in 1863	350	0.123	0.329	0	1
<b>Main controls</b>					
Land suitability for cereals	350	16.678	11.025	1.490	66.380
Land suitability for citrus	350	15.043	7.717	0	48
Land suitability for olive	350	29.108	12.943	0	69.273
Active sulphur cave in 1868-1870	350	0.106	0.308	0	1
Banditry in 1820-1849	350	0.249	0.433	0	1
<b>Geographic controls</b>					
Underground water presence	350	0.631	0.483	0	1
Latitude	350	37.719	0.346	36.715	38.232
Longitude	350	14.318	0.797	12.434	15.554
Average altitude (in meters)	350	547.247	332.022	24	1,932
Altitude of town center (in meters)	350	411.212	278.910	2	1,212
Distance from non-seasonal river (in km)	350	7.503	7.114	0.012	33.555
Distance from commercial port (in km)	350	37.815	20.325	0	86.436
Distance from post road (in km)	350	3.855	4.172	0	21.211
Distance from regional capital (in km)	350	162.840	89.188	0	357.798
Distance from department capital (in km)	350	23.434	13.628	0	62.355
<b>State presence controls</b>					
Ln population density in 1861	350	4.829	1.053	1.914	12.922
Elementary schools in 1862 per 1,000 people	350	0.571	0.411	0	2.809
Electoral turnout in 1861	350	0.838	0.134	0	1
Archbishopric or bishopric in 1839	350	0.037	0.189	0	1
Garibaldi in 1860	350	0.454	0.499	0	1

*Notes:* See Section 3 for variable definitions and sources.

Table A2: Means Comparison

	Full sample			Restricted sample		
	<i>Repr</i> 1863 = 0 (1)	<i>Repr</i> 1863 = 1 (2)	Difference (3) = (1)-(2)	<i>Repr</i> 1863 = 0 (4)	<i>Repr</i> 1863 = 1 (5)	Difference (6) = (4)-(5)
<b>Main controls</b>						
Land suitability for citrus	14.75	17.15	-2.40**	14.75	16.51	-1.75
Land suitability for olive	28.70	32.02	-3.32**	28.53	31.82	-3.28**
Land suitability for cereals	16.08	20.95	-4.87**	16.01	19.38	-3.38*
Active sulphur cave in 1868-1870	0.09	0.19	-0.09	0.09	0.19	-0.10
Banditry in 1820-1849	0.20	0.60	-0.41***	0.19	0.67	-0.47***
<b>Geographic controls</b>						
Underground water presence	0.67	0.35	0.32***	0.67	0.39	0.29***
Latitude	37.71	37.81	-0.10**	37.71	37.80	-0.09*
Longitude	14.45	13.41	1.04***	14.45	13.45	0.99***
Average altitude (in meters)	551.99	513.38	38.61	551.33	525.33	26.00
Altitude of town center (in meters)	414.98	384.32	30.66	413.55	426.59	-13.04
Distance from non-seasonal river (in km)	7.82	5.27	2.54**	7.85	4.03	3.82***
Distance from commercial port (in km)	38.82	30.66	8.15**	38.54	32.40	6.15*
Distance from post road (in km)	4.06	2.39	1.67**	4.10	2.86	1.24*
Distance from regional capital (in km)	175.42	73.01	102.42***	175.61	71.94	103.67***
Distance from department capital (in km)	24.27	17.47	6.80***	24.39	19.55	4.84**
<b>State presence controls</b>						
Ln population density in 1861	4.83	4.86	-0.03	4.83	4.84	-0.01
Elementary schools in 1862 per 1,000 people	0.60	0.39	0.20***	0.60	0.40	0.20***
Electoral turnout in 1861	0.85	0.78	0.07***	0.85	0.78	0.06**
Archbishopric or bishopric in 1839	0.03	0.12	-0.09*	0.02	0.06	-0.03
Garibaldi in 1860	0.40	0.86	-0.46***	0.39	0.83	-0.44***
N	307	43	350	304	36	340

*Notes:* Columns 1 and 2 report, respectively, means for towns that were not hit by the 1863 repression campaign (control) and for towns that instead were (treatment), in the full sample. Columns 4 and 5 report means for the same groups, but in the restricted sample, which excludes the ten garrison towns discussed in Section 4.3. Columns 3 and 6 report differences of group means between, respectively, Columns 1 and 2 and Columns 4 and 5. The significance test for the difference between means allows for unequal variance. \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A3: Determinants of Government Repression

Dependent variable:	ln(odds of Repression 1863 = 1)			
	Full sample		Restricted sample	
	Coeff.	S.E.	Coeff.	S.E.
	(1)	(2)	(3)	(4)
<b>Main controls</b>				
Land suitability for citrus	-0.068	(0.090)	-0.114	(0.121)
Land suitability for olive	-0.093**	(0.043)	-0.084*	(0.047)
Land suitability for cereals	-0.011	(0.051)	0.024	(0.046)
Active sulphur cave in 1868-1870	2.106***	(0.586)	2.181***	(0.599)
Banditry in 1820-1849	-0.152	(0.529)	-0.076	(0.562)
<b>Geographic controls</b>				
Underground water presence	-1.330***	(0.495)	-1.099**	(0.437)
Latitude	1.259	(1.569)	1.243	(1.542)
Longitude	-3.249***	(0.875)	-2.922***	(0.576)
Average altitude (in meters)	0.000	(0.001)	0.000	(0.001)
Altitude of town center (in meters)	-0.000	(0.001)	0.000	(0.001)
Distance from non-seasonal river (in km)	-0.112**	(0.049)	-0.216***	(0.055)
Distance from commercial port (in km)	0.007	(0.018)	0.005	(0.020)
Distance from post road (in km)	-0.139**	(0.066)	-0.139**	(0.059)
Distance from regional capital (in km)	-0.002	(0.010)	-0.004	(0.009)
Distance from department capital (in km)	-0.020	(0.020)	-0.022	(0.019)
<b>State presence controls</b>				
Ln population density in 1861	-0.205	(0.174)	-0.169	(0.175)
Elementary schools in 1862 per 1,000 people	-2.631***	(0.727)	-2.951***	(0.829)
Electoral turnout in 1861	-2.212**	(0.866)	-2.197**	(0.915)
Archbishopric or bishopric in 1839	1.455	(1.582)	-0.392	(1.292)
Garibaldi in 1860	1.618**	(0.677)	1.572**	(0.690)
N	350		340	
McFadden's R <sup>2</sup>	0.493		0.474	

*Notes:* Maximum Likelihood estimates of the impact of the control variables used in the main analysis on the exposure to government repression in 1863, in the full sample (Columns 1 and 2) and in the restricted sample (Columns 3 and 4), which excludes the ten garrison towns discussed in Section 4.3. Standard errors clustered by department are reported in parentheses. \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A4: Baseline Estimates with Department FE

Dependent variable:	Mafia 1875					
	(1)	(2)	(3)	(4)	(5)	(6)
Repression 1863	0.443*** (0.098)	0.338*** (0.078)	0.332*** (0.076)	0.316*** (0.079)	0.287*** (0.092)	0.271*** (0.091)
Department FE	No	Yes	Yes	Yes	Yes	Yes
Main controls except land concentration	No	No	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes	Yes
State presence controls	No	No	No	No	Yes	Yes
High and medium land concentration	No	No	No	No	No	Yes
N	350	350	350	350	350	294
No. of clusters	24	24	24	24	24	24
P-value of wild cluster bootstrap test	0.000	0.002	0.002	0.002	0.002	0.000
Mean of dependent variable	0.123	0.123	0.123	0.123	0.123	0.133
R <sup>2</sup>	0.196	0.342	0.359	0.404	0.429	0.465
Oster (2019)'s $\delta$		1.728	1.680	1.637	1.477	1.339

*Notes:* OLS estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1875 at the town level. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. In column (6) the dummy *Low land concentration* is omitted from the model. Standard errors clustered by department are reported in parentheses. The p-value of the wild cluster bootstrap test of the hypothesis  $Repr1863 = 0$  is reported (bootstrap replications = 999). Oster (2019)'s relative degree of selection  $\delta$  is estimated using  $R_{max} = 1.3R^2$ . \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A5: Baseline Estimates with Spatial HAC Standard Errors

Dependent variable:	Mafia 1875					
	(1)	(2)	(3)	(4)	(5)	(6)
Repression 1863	0.443***	0.338***	0.344***	0.314***	0.274***	0.252***
[15 km cut-off]	(0.088)	(0.079)	(0.082)	(0.071)	(0.073)	(0.068)
[30 km cut-off]	(0.092)	(0.074)	(0.074)	(0.063)	(0.086)	(0.069)
[0.4 dd latitude, 0.75 dd longitude cut-off]	(0.093)	(0.082)	(0.086)	(0.084)	(0.091)	(0.086)
Province FE	No	Yes	Yes	Yes	Yes	Yes
Main controls except land concentration	No	No	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes	Yes
State presence controls	No	No	No	No	Yes	Yes
High and medium land concentration	No	No	No	No	No	Yes
N	350	350	350	350	350	294
R <sup>2</sup>	0.196	0.275	0.293	0.333	0.367	0.401

*Notes:* OLS estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1875 at the town level. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. In column (6) the dummy *Low land concentration* is omitted from the model. Conley's spatially-corrected standard errors are reported in parentheses. In the first two cases I allow errors to be correlated within a circle of radius 15 or 30 kilometers around each town, following the methodology introduced by Colella et al. (2019). In the third case I employ a cutoff of 0.4 decimal degree for latitude and 0.75 decimal degree for longitude, which represent roughly half of the height and width of the Sicilian island (Acemoglu et al., 2020). \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A6: Baseline Estimates using a Conditional Logit Model

Dependent variable:	ln(odds of Mafia 1875 = 1)					
	(1)	(2)	(3)	(4)	(5)	(6)
Repression 1863	2.658*** (0.490)	1.847*** (0.378)	1.848*** (0.429)	2.084*** (0.585)	1.548** (0.634)	2.246** (0.926)
Department FE	No	Yes	Yes	Yes	Yes	Yes
Main controls except land concentration	No	No	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes	Yes
State presence controls	No	No	No	No	Yes	Yes
High and medium land concentration	No	No	No	No	No	Yes
N	350	169	169	169	169	147
No. of clusters	24	12	12	12	12	12
P-value of score cluster bootstrap test	0.000	0.001	0.002	0.000	0.009	0.000
McFadden's R <sup>2</sup>	0.184	0.120	0.152	0.320	0.453	0.572

*Notes:* Maximum Likelihood estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1875 at the town level, obtained from a conditional logit model. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. In column (6) the dummy *Low land concentration* is omitted from the model. Standard errors clustered by department are reported in parentheses. The p-value of the score cluster bootstrap test of the hypothesis  $Repr1863 = 0$  is reported (bootstrap replications = 999). \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A7: Nearest Neighbour Matching Estimates of ATET

Dependent variable:	Mafia 1875					
	Matching characteristics: full set of controls			Matching characteristics: latitude and longitude		
	(1)	(2)	(3)	(4)	(5)	(6)
Repression 1863	0.202* (0.108)	0.144 (0.108)	0.205** (0.091)	0.206* (0.110)	0.284*** (0.088)	0.323*** (0.086)
Main controls	Yes	Yes	Yes	No	No	No
Geographic controls	Yes	Yes	Yes	No	No	No
State presence controls	Yes	Yes	Yes	No	No	No
Latitude and longitude	No	No	No	Yes	Yes	Yes
N	350	350	350	350	350	350
No. of matches	1	2	4	1	2	4
Mean distance (km)				7.7	9.4	12.1
Median distance (km)				7.1	8.6	10.5

*Notes:* OLS estimates. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The treatment variable is a dummy denoting exposure to government repression in 1863 and the average treatment effect for the treated (ATET) is reported. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. The matching estimator is corrected for a large-sample bias that exists when matching on more than one continuous covariate (Abadie and Imbens, 2006; 2011). Heteroskedasticity-robust standard errors adjusted for the requested number of matches are reported. Distance between each repressed town and its  $k$  closest non-repressed matches is finally reported, with  $k = 1, 2$  or  $4$ . \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.



Table A8: IV Analysis with Alternative Distance Transformations

	(1)	(2)	(3)	(4)	(5)
<i>Excluded Instrument: IHS(Distance)</i>					
Panel A: Second stage	Dependent variable: Mafia 1875				
Regression 1863	0.809*** (0.156)	0.835*** (0.206)	0.877*** (0.230)	0.811*** (0.211)	0.789*** (0.216)
Mean of dependent variable	0.106	0.106	0.106	0.106	0.106
Panel B: First stage	Dependent variable: Repression 1863				
IHS(Distance from least-cost path)	-0.094*** (0.014)	-0.084*** (0.021)	-0.081*** (0.022)	-0.085*** (0.022)	-0.085*** (0.022)
Kleibergen-Paap Wald F statistic	44.11	15.41	13.46	14.61	14.68
<i>Excluded Instrument: Distance dummy</i>					
Panel C: Second stage	Dependent variable: Mafia 1875				
Regression 1863	0.795*** (0.169)	0.743*** (0.233)	0.841*** (0.259)	0.832*** (0.251)	0.785*** (0.240)
Mean of dependent variable	0.106	0.106	0.106	0.106	0.106
Panel D: First stage	Dependent variable: Repression 1863				
$\mathbb{1}\{\text{Distance from least-cost path} < 5\text{km}\}$	0.337*** (0.075)	0.207*** (0.078)	0.206*** (0.078)	0.200** (0.078)	0.212*** (0.075)
Kleibergen-Paap Wald F statistic	20.15	7.11	7.05	6.50	8.03
Province FE	No	Yes	Yes	Yes	Yes
Main controls	No	No	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes
State presence controls	No	No	No	No	Yes
N	340	340	340	340	340
No. of clusters	24	24	24	24	24

*Notes:* 2SLS (Panels A and C) estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1875 at the town level. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. This regressor is instrumented by, alternatively, the inverse hyperbolic sine transformation of the distance to the network of least-cost paths described in Section 4.3 (Panels A and B), or a dummy equal to 1 if a town falls within 5 kilometers of the aforementioned network, and equal to 0 otherwise (Panels C and D). Panels B and C report the results of the first-stage regressions for the exposure to government repression, using the two excluded instruments just mentioned. The ten garrison towns used to construct the optimal route network are dropped from the sample. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. Standard errors clustered by department are reported in parentheses. \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A9: IV Analysis with Spatial HAC Standard Errors

Dependent variable:	Mafia 1875				
	(1)	(2)	(3)	(4)	(5)
Repression 1863	0.812***	0.851***	0.889***	0.809***	0.788***
[15 km cut-off]	(0.171)	(0.266)	(0.291)	(0.172)	(0.184)
[30 km cut-off]	(0.179)	(0.222)	(0.256)	(0.170)	(0.185)
Province FE	No	Yes	Yes	Yes	Yes
Main controls	No	No	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes
State presence controls	No	No	No	No	Yes
N	340	340	340	340	340
Kleibergen-Paap Wald F statistic	36.71	16	12.89	16.47	16.31

*Notes:* 2SLS estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1875 at the town level. The dependent variable is a dummy indicating the presence of mafia in the town, based on information included in the Borsani-Bonfadini parliamentary enquiry (Carbone and Grispo, 1969). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. This regressor is instrumented by the natural logarithm of (one plus) the distance to the network of least-cost paths described in Section 4.3. The ten garrison towns used to construct this optimal route network are dropped from the sample. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. Conley's spatially-corrected standard errors are reported in parentheses. I allow errors to be correlated within a circle of radius 15 or 30 kilometers around each town, following the methodology introduced by Colella et al. (2019). \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A10: Persistence Analysis using a Ordered Logit Model

Dependent variable:	Mafia 1900				
	(1)	(2)	(3)	(4)	(5)
Repression 1863	1.832*** (0.315)	1.023*** (0.316)	0.785*** (0.303)	0.845** (0.382)	0.741* (0.380)
Department FE	No	Yes	Yes	Yes	Yes
Main controls	No	No	Yes	Yes	Yes
Geographic controls	No	No	No	Yes	Yes
State presence controls	No	No	No	No	Yes
N	277	272	272	272	272
No. of clusters	24	23	23	23	23
McFadden's R <sup>2</sup>	0.041	0.020	0.087	0.148	0.176

*Notes:* Maximum Likelihood estimates of the impact of exposure to government repression in 1863 on the presence of mafia in 1900 at the town level, obtained from an ordered logit model run in the restricted sample excluding the ten garrison towns discussed in Section 4.3. The dependent variable is an ordinal variable indicating the level of mafia activity in 1900 on a 0 to 3 scale (0 meaning no mafia presence, 3 meaning high mafia presence), according to Cutrera (1900). The main explanatory variable is a dummy equal to 1 if government repression occurred in 1863 in the town, and equal to 0 otherwise. Main controls include: land suitability for citrus, olives and cereals, dummy for the presence of sulphur mines in 1868-1870 and dummy for banditry presence in 1820-1849. Geographic controls include: dummy for underground water presence, latitude, longitude, average altitude, altitude of town centre, distance from closest non-seasonal river, distance from closest commercial port, distance from closest post road, distance from regional capital Palermo and distance from department capital. State presence controls include: log of population density in 1861, number of elementary schools in 1862 (over 1861 population), electoral turnout in 1861, dummy for presence of bishopric or archbishopric in 1839 and dummy for presence of *garibaldini* in 1860. Standard errors clustered by department are reported in parentheses. \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.

Table A11: Event Study Analysis

Dependent variable:	Draft evasion rate
Repression share 1863 × 1860	0.036 (0.037)
Repression share 1863 × 1861	0.060 (0.036)
Repression share 1863 × 1863	0.004 (0.022)
Repression share 1863 × 1864	-0.036* (0.021)
Repression share 1863 × 1865	-0.050** (0.020)
Repression share 1863 × 1866	-0.027* (0.014)
Repression share 1863 × 1867	-0.039** (0.017)
Repression share 1863 × 1868	-0.025** (0.010)
Department FE	Yes
Year FE	Yes
Number of departments	24
N	384
R <sup>2</sup>	0.384
Within R <sup>2</sup>	0.699
Mean dependent variable 1862	0.246

*Notes:* OLS estimates of the impact of exposure to government repression in 1863 on the draft evasion rate across Sicilian departments, from the event study model in Equation (4). Standard errors clustered by department are reported in parentheses. \* = p-value < 10%, \*\* = p-value < 5%, \*\*\* = p-value < 1%.