Mafias and Firms*

Jaime Arellano-Bover† Marco De Simoni‡ Luigi Guiso§ Rocco Macchiavello¶ Domenico J. Marchetti|| Mounu Prem**

January 2025

Abstract

Infiltration of legitimate businesses by criminal organizations (OCGs) is potentially significant. Despite the policy relevance, its extent and underlying motives remain uncertain. We propose a framework to distinguish OCGs' motives and validate it using new data from the Italian Financial Intelligence Unit. About 2% of Italian firms appear to have links with OCGs. We depart from the dominant idea in the literature that infiltrated firms are always *contaminated* with criminal activities. While this characterization applies well to smaller and medium-sized firms, often directly established by OCGs, many firms are already well-established when infiltrated. OCGs separate these firms from criminal activities to pursue pecuniary and non-pecuniary benefits, such as political connections. This previously undetected motive – labeled *pure*, in contrast to traditional *contaminated* ones – accounts for a substantial share of OCG infiltration, nuances our understanding of infiltration's relationship to money laundering, and yields novel policy implications.

*This project would not have been possible without the collaboration of Luca Baron and Flavio Ferlanti at UIF. We benefited from a workshop with prosecutors at the *Direzione Nazionale Antimafia e Antiterrorismo* and officers at *Guardia di Finanza* and *Direzione Investigativa Antimafia*. We also thank Marianne Bertrand, Jordi Blanes-i-Vidal, Francois Bourguignon, Lorenzo Casaburi, Emanuele Colonnelli, Ray Fisman, Sauro Mocetti, Elias Papaioannou, Jason Sherman, Alfredo Tidu, David Yanagizawa, and seminar and conference participants at Al Capone, Bank of Italy, Bahamas AML Conference, Berkeley, Bocconi, CEPR IMO-ENT, EBRD-IFC, HEC Lausanne, IADB, IFS-STICERD, LSE, PSE, Yale and Zurich for useful comments. Laura Pittalis and Oliver Seager provided excellent research assistance. This project uses highly confidential data from the UIF's *Mappatura*. The *Mappatura* identifies "firms potentially connected to organized crime" (UIF (2021), pp. 46-47). For simplicity, we refer to such firms as "infiltrated". The views expressed in this paper are those of the authors and do not involve the responsibility of UIF or the Bank of Italy. The usual disclaimers apply.

†Yale. email: jaime.arellano-bover@yale.edu.

‡UIF, Bank of Italy. email: marco.desimoni@bancaditalia.it.

§EIEF. email: luigi.guiso55@gmail.com.

¶LSE. email: r.macchiavello@lse.ac.uk.

||UIF, Bank of Italy. email: domenico.marchetti@bancaditalia.it.

**EIEF. email: francisco.munoz@eief.it.

1 Introduction

Proceeds from organized criminal activities are estimated at USD 2.1 trillion, or 3.6% of global GDP (UNODC, 2011). Economic logic suggests that organized crime groups (OCGs) face limits to the reinvestment of criminal profits into the criminal industry.[1] This implies that the share of the legal economy infiltrated by OCGs is potentially substantial, though how large remains to be established. Perhaps more importantly than its volume, the *motives* of infiltration are also a matter of interest as they may affect the type of policies necessary to deter organized crime. The relevance notwithstanding, serious challenges have hindered empirical progress. Besides well-known data limitations (OCGs' activities are notoriously difficult to measure), the motives of OCGs' infiltration are not observable and must be inferred from observable behavior. This requires a conceptual framework tailored to the available data.

This paper makes progress in providing new, more reliable evidence on the pervasiveness of OCGs' infiltration into the legal economy and, crucially, examines the motives behind such infiltration. We develop a new conceptual framework for OCGs' infiltration of legitimate firms. The model delivers a rich set of testable predictions that allow us to infer the underlying motives of infiltration from the firm's observable behavior. We apply the framework to Italy, where the endemic presence of OCGs and a sophisticated legal and institutional apparatus to fight them, provide an ideal setting for our analysis. In particular, we leverage the *Mappatura*, a new and highly confidential dataset assembled, and used, by the Financial Intelligence Unit of the Bank of Italy (UIF) to map OCGs' infiltration in legitimate firms. Over our sample period, the *Mappatura* identifies around 100,000 firms as potentially connected to OCGs. Although, by its nature, we can never have a perfectly accurate map of OCGs' infiltration, the *Mappatura* arguably represents the most comprehensive attempt ever undertaken. This large sample of potentially infiltrated firms is also essential to uncover the broad set of motives identified by our conceptual framework.

Our main contribution is to unveil a previously unnoticed motive of infiltration and to show that it accounts for a sizable share of OCGs' infiltration. Traditionally, both the legislator and the existing literature have assumed that the infiltrated firm benefits from criminal activities (e.g., corruption in public

[1] For example, OCGs cannot advertise to acquire new customers without increasing the risk of detection and must rely on a few trusted intermediaries to conduct their criminal activities.

procurement) or is used for criminal activities (e.g., money laundering).[2] In both cases, the firm is directly "contaminated" by OCGs' criminal activities. Alongside these traditional motives of infiltration, our conceptual framework highlights – and the analysis of the *Mappatura* supports – a novel motive in which the infiltrated firm is *not* directly involved in criminal activities. This type of infiltration – which we label "pure" to distinguish it from the various "contaminated" ones – yields pecuniary and non-pecuniary benefits, such as relationships with entrepreneurs and politicians, that are extremely valuable to the OCG.[3] As we will discuss, this new perspective on OCGs' infiltration of legitimate firms has important policy implications.

Section 2 presents the *Mappatura*: its construction, advantages, and limitations. In a nutshell, the *Mappatura* starts with a highly confidential list of individuals who are under investigation for, or are reported in judicial documents connected to, OCGs-related crimes *and* are also involved in suspicious financial transactions reported to UIF. Following the literature, we define infiltration of an OCG as "any case in which a natural person belonging to a criminal organization or acting on its behalf, [...], invests financial and/or human resources to participate in the decision-making process of a legitimate business" (Transcrime, 2017, p.19). The key feature of this definition is that a person tied to an OCG plays an active role in the decision process of the firm. Hence, in our baseline definition, a firm is infiltrated if it has at least an owner *or* an administrator on the highly confidential list of individuals described above.

Section 3 describes the more than 100,000 firms identified in the *Mappatura* over our sample period. These firms represent around 2% of all corporations and partnerships in Italy. Although the incidence of infiltrated firms is higher in the South, where the most important Italian OCGs originate, the majority of infiltrated firms are located in the economically more prosperous Northern regions. The firms identified in the *Mappatura* are larger than the typical Italian firm and account for around a tenth of the aggregate employment and revenues of the Italian private sector. It is worth emphasizing that the firms included in the *Mappatura* cannot be deemed with certainty to be infiltrated or controlled by or connected to organized crime, a circumstance that can only be ascertained

[2]For example, the 1967 Taskforce on Organized Crime - which led to the U.S. Racketeer Influenced and Corrupt Organizations (RICO) Act - defines the provision of illicit goods and services as OCGs core activities, whereas when OCGs "turn to legitimate business, they terrorize, blackmail, and monopolize." (Schelling, 1971, p. 180). Article 416-bis of the Italian Penal Code defines the infiltrated firm as one that benefits from the intimidation force of the OCG to acquire economic benefits.

[3]Of course, there is nothing pure about this motive: the individuals involved are criminals.

at the end of a judicial procedure. Still, these figures suggest that the infiltration of legitimate firms by OCGs is a pervasive phenomenon in Italy.

Section 4 presents our conceptual framework. An OCG deciding whether to contaminate the infiltrated firm with criminal activities or not faces a trade-off. The OCG can exploit the firm to conduct criminal activities (what we label the *functional* motive); or it can use criminal activities to boost the firm's performance (the *competitive* motive). In both cases, the OCG benefits from directly contaminating the firm's operations with its criminal activities. This benefit, however, comes at the cost of a higher risk of detection and confiscation. There is thus an alternative – which we label *pure* motive – in which the OCG keeps the firm separate from criminal activities. In this case, the OCG obtains pecuniary and non-pecuniary benefits from infiltration. For simplicity, the model focuses on a pecuniary benefit: the infiltrated firm yields a higher return than alternatives to the OCG's funds.[4]

This trade-off delivers a rich set of predictions. First, the infiltration motives can be distinguished in the data. Relative to an otherwise identical clean firm, the *functional* and *competitive* motives distort the infiltrated firm's operation, while the *pure* motive does not. Internal funds substitute for external ones in the *pure* motive, but not necessarily in the two other motives. Second, the model characterizes the type of infiltration: small firms are used to conduct illegal activities (*functional* motive), medium-sized firms benefit from criminal activities (*competitive* motive), the largest firms – for which a higher risk of detection and confiscation is particularly costly – are kept separate from criminal activities (the *pure* motive).

Section 5 tests the model's predictions and infers the relative prevalence of the different motives. Firms in the *Mappatura* are roughly split in half between *born-infiltrated* firms, in which the individual tied to the OCG is present at birth, and *born-clean* ones, in which the individual connected to an OCG enters at a later date. This distinction is appealing both on conceptual and empirical grounds. On the conceptual front, *born-infiltrated* firms are smaller than *born-clean* firms. According to our model, they are more likely to reflect the riskier *contaminated* (*functional* or *competitive*) motives, while *born-clean* firms the *pure* one. Consistent with this, and relative to born-clean firms, born-infiltrated firms are significantly more likely to have been confiscated by judicial authorities and more prevalent in regions with weaker institutions, home of the main OCGs. This gives us some confidence that the distinction between born-clean

[4]We postpone a discussion of, and evidence on, non-pecuniary benefits to Section 6.

and born-infiltrated firms captures different underlying motives. On the empirical front, the two groups differ in the strategies available to construct a suitable comparison group of non-infiltrated firms. We thus organize the empirical analysis separating *born-clean* and *born-infiltrated* firms.

We consider *born-clean* firms first and – as predicted by the model – uncover the *pure* motive's hallmarks: no change in the firm's operation and substitution away from bank loans. For born-clean firms, it is possible to explore changes in outcomes around the time of infiltration within a staggered DID framework. Because infiltration occurs, by definition, at the same time as a change in at least an owner or an administrator, we compare firms that become infiltrated to those that experience a similar change – a "clean inflow". On average, infiltration is associated with changes in sales, employment, and input purchases identical to those of firms experiencing clean inflows. However, while non-infiltrated firms attract *more* bank loans after a clean inflow, infiltration is associated with *less* bank loans, and a contemporaneous increase in liquidity.[5] In line with the model, the average behavior masks significant heterogeneity: infiltration in smaller *born-clean* firms is associated with an expansion in the firm's operations, consistent with the *competitive* motive.

Turning to the smaller, born-infiltrated firms, we also confirm the model predictions and find the hallmarks of the *functional* motive: an inflated scale of operation, associated with worse economic performance. Unlike born-clean firms, born-infiltrated firms are, by definition, not observed before infiltration. We thus compare them to firms established in the same year, province, and sector. Across all operational outcomes – revenues, employment, and inputs – born-infiltrated firms are initially larger, but subsequently grow at a slower pace and are less profitable than firms in the comparison group. These patterns starkly contrast with the findings for born-clean firms, highlighting the likely different motives of infiltration between the two groups of firms.

The evidence supports our distinction between the *contaminated* and *pure* motives of infiltration. This distinction departs from the dominant idea in the literature that infiltration is always contaminated with criminal activities. Such characterization applies well to smaller and medium-sized firms, often directly established by the OCGs. Many firms, however, are already large and well-established when infiltration occurs. For these firms, which account for 85 per-

[5] Unlike clean inflows, infiltration is preceded by a decrease in liquidity, suggesting that OCGs tend to target firms suffering a liquidity shortage. Once infiltration occurs, however, liquidity is so abundant that loans from banks – which are more intrusive and possibly more expensive – become less appealing and are discontinued or less utilized.

cent of the assets of the firms in the *Mappatura*, infiltration mostly reflects the *pure* motive, in which the OCG keeps the firm separate from criminal activities in the pursuit of other pecuniary and non-pecuniary benefits. This previously undetected *pure* motive is thus a significant, if not the predominant, motive of infiltration. Section 6 concludes by discussing non-pecuniary benefits likely associated with it, its relationship with money laundering, and the resulting policy implications.

We conjecture that the acquisition of what we label "relational capital" – a web of relationships with important actors in the legal economy (such as board members of other large firms, industry associations, public administration, politicians, consultants, etc.) – is a key non-pecuniary benefit of *pure* infiltration. This benefit can be accessed almost exclusively, if not *only*, through competent, seemingly respectable, individuals trusted by the OCG and directly involved in the operations of large, legitimate, businesses. This conjecture naturally leads to the hypothesis that infiltrat*ed* firms and infiltrat*ing* individuals differ in the extent and composition of their connections. We find strong empirical support. Infiltration, particularly among large, born-clean firms, targets firms with board members more connected to other firms and more likely to be elected politicians at the municipal, regional, national, and European levels. Infiltrating individuals are also more connected to other firms, even compared to other board members of infiltrated firms.

The *pure* motive is not money laundering (ML), which is a crime, and thus belongs to the *functional* motive. ML converts the proceeds of crime into assets with a legitimate appearance. The criminology literature, primarily based on investigations and court cases, often finds evidence limited to the *placement* and *layering* stages of ML schemes, where funds are introduced into the financial system and their origins concealed (Gilmore, 2004; Riccardi and Reuter, 2024). Legitimate firms participate in these stages of ML schemes through false invoicing, implying anomalies in revenues and/or input purchases. We find no such anomalies among large, born-clean firms. *Pure* infiltration is thus not used for *placement* and *layering*, but provides a rare glimpse into the final, and most elusive, *integration* stage in which laundered funds are deployed.

The empirical relevance of the *pure* motive of infiltration has policy implications. The optimal allocation of scarce resources to fight organized crime and the design of monitoring and leniency programs and screening algorithms depend on whether OCGs involve legitimate businesses in criminal activities or not. But a more concerning implication emerges considering our evidence that the *pure* motive hides OCGs' desire, and ability, to interact with the main players of the legal economy – including politicians. In due time, these connections can become political power and, through lobbying (Bertrand et al., 2014, Bertrand et al., 2023), influence policymaking (e.g., anti-money-laundering and financial regulation) thus strengthening and perpetuating OCGs' grip on the economy and society at large.

Related Literature This paper contributes to several strands of the literature. Organized crime is a pervasive phenomenon, particularly in Latin America, Eastern Europe, Asia, and Western Africa (Pinotti, 2015a). Accordingly, recent contributions have studied OCGs in different contexts. For example, in Colombia, Blattman et al. (2024) study the relationship between state presence and gangs; in El Salvador, the rise of gangs stifled economic development (Melnikov et al., 2020) with the costs of their extortion passed through to consumer prices (Brown et al., 2024); in Nigeria's oil industry connections to OCGs give local producers an advantage relative to foreign companies which are exposed to violence and thefts (Rexer, 2022). Our novel conceptual framework can shed light on the infiltration of legitimate businesses across these diverse contexts.

OCGs increasingly pose serious threats in advanced economies as well. (Transcrime, 2017) reports 2,380 references to OCGs' infiltration of firms across five European Countries (the UK, the Netherlands, Italy, Sweden, and Slovenia). We focus on Italy, home to some of the oldest OCGs worldwide and a notable exception among higher-income countries.[6] Italian OCGs have stifled the socio-economic and political development of the country. On the economic front, Pinotti (2015b) estimates that OCGs' presence lowered regional GDP per capita by 16%. The dismissal of city councils infiltrated by organized crime increases employment, the number of firms, industrial real estate prices (Fenizia and Saggio, 2023), and competition in procurement (Slutzky and Zeume, 2024). We contribute novel facts on the overall incidence of firms potentially connected to OCGs in the economy and their diffusion across regions and sectors. On the political front, Italian OCGs have historically used violence to influence elections (Alesina et al., 2019) and curb political competition (Acemoglu et al., 2020). Our analysis hints at a less visible channel: the infiltration of legitimate firms to acquire political connections.

Our paper is more directly related to the – primarily empirical – literature on

[6] The main Italian OCGs – the Sicilian Mafia, the Camorra, and the 'Ndrangheta – emerged in Southern Italy during the 19^{th} century (Gambetta (1996), Lupo (2009), Bandiera (2003), Buonanno et al. (2015), Dimico et al. (2017)) and then expanded to other regions (Varese, 2006).

OCGs' infiltration of legitimate firms. Several papers have extended the canonical Beckerian model of crime to study criminal organizations (e.g., Buchanan, 1973, Backhaus, 1979, Fiorentini and Peltzman, 1997, and Dixit, 2004), but we are not aware of models of OCGs' motives to infiltrate firms. The criminology literature has developed taxonomies of infiltrated firms (see, e.g., Arlacchi, 2010, Parbonetti, 2021) that are, however, not suited to our purpose. First, we need testable hypotheses to infer the motive of infiltration from observed behavior. More importantly, these taxonomies are developed from the behavior of firms caught in criminal investigations and therefore *assume* that the firm is involved in criminal activities. In contrast, our conceptual framework and empirical evidence highlight that this is not necessarily the case.

On the empirical front, Le Moglie and Sorrenti (2022) find that provinces with a higher presence of OCGs experienced a milder reduction in firms' entry after the 2008 financial crisis, suggesting that OCGs helped firms overcome the credit crunch. Daniele and Dipoppa (2023) document the strategic behavior of firms participating in public procurement projects to elude screenings to detect OCG-connected firms. Calamunci and Drago (2020) find that the assignment of infiltrated firms to judicial investigations has a positive spillover on competing firms, suggesting a burden imposed by infiltrated firms on other firms.

Mirenda et al. (2022) study the effect of 'Ndrangheta infiltration on firms' performance and is most closely related to ours. They propose a creative approach in which infiltration is proxied by whether the firm's owners and/or directors have family names associated with OCGs. Focusing on *born-clean* firms in a DID framework, they find that infiltration generates a significant rise in firms' revenues followed by exit, and argue that OCGs use infiltration predominantly for money laundering and/or rent extraction (our *functional* motive). Our analysis differs from theirs in important ways and paints a different picture of OCGs' infiltration in Italy, with radically different policy implications. We model – and find evidence consistent with – a wider set of motives for infiltration. Alongside the *functional* motive (which we however detect on born-infiltrated firms), we find evidence for a novel *pure* motive on *born-clean* firms. Not only restricting attention to *born-clean* firms paints a partial picture of OCGs' infiltration motives but also controlling for clean inflows in the DID – as we do – reverses Mirenda et al. (2022) conclusions, even when using their infiltration definition and data.

2 Background and Data

This section introduces Italy's main OCGs (Section 2.1) and the *Mappatura* (Section 2.2). Appendix A describes additional data sources.

2.1 OCGs in Italy

Italy has a pervasive presence of autochthonous OCGs, providing a natural canvas to study OCGs' infiltration in the legal economy. The main OCGs in Italy are the Sicilian Mafia, the Camorra, and the 'Ndrangheta. Originating from their respective regions (Sicily, Campania, and Calabria, all in the South of Italy), these OCGs have expanded into other regions in the traditionally richer Northern part of the country, as well as abroad. They dominate illicit activities but also infiltrate the legal economy, posing a significant challenge to law enforcement and governance. Italy has thus developed a comprehensive regulatory framework aimed at countering them (DNA, 2020).

The Sicilian Mafia, whose origins can be traced back to the 19th century (see Gambetta, 1996 and Lupo, 2009), is perhaps the most widely known, at least in part because of its historical connections to OCGs in the U.S. (see, e.g., Mastrobuoni and Patacchini, 2012 and Mastrobuoni, 2015). The Sicilian Mafia is characterized by a centralized hierarchy where a committee controls and coordinates criminal families. A landmark trial, the so-called Maxiprocesso, convicted numerous members of the central committee. It is now generally believed that the organization's influence has been dented.[7] In contrast to the Sicilian Mafia, the Camorra, is characterized by smaller clans often in fierce competition with each other.

The 'Ndrangheta – which also originated in the 19th century in the southern region of Calabria but then expanded nationwide and abroad (Varese, 2006, Ciconte, 2008) – is organized around tightly closed family-based clans. In contrast to what was generally believed, recent investigations have demonstrated that the 'Ndrangheta also has a centralized committee that coordinates the activities of the different clans, helps form alliances to undertake large-scale illegal activities, and settles disputes. The family-based structure and the secrecy of the highest layer of the organization (itself unknown to lower-level members) have made it difficult to counter 'Ndrangheta. Notwithstanding notable law enforce-

[7]The aftermath of the Maxiprocesso reached a tragic climax with the assassinations of two prominent anti-Mafia prosecutors, Giovanni Falcone and Paolo Borsellino, in 1992, and subsequent terrorist attacks.

ment efforts and successes in recent years, the 'Ndrangheta is among the richest and most powerful global OCGs (Europol, 2013) with revenues from illicit activities estimated in 2010 at over 3.5 billion euro, nearly twice as much as those of the Sicilian Mafia (Transcrime, 2015). Of these revenues, only a quarter are estimated to be produced in the organization's region of origin, in contrast to the two-thirds estimated for the Sicilian Mafia and the Neapolitan Camorra. Although we cannot distinguish infiltration by the different OCGs, most of the infiltration in the *Mappatura* is likely tied to the 'Ndrangheta.

2.2 OCGs Infiltration in the Legal Economy: The *Mappatura*

The Construction of the *Mappatura* This project leverages the *Mappatura* – UIF's most systematic effort to map OCGs infiltration in the Italian legal economy to date. We follow the literature and define OCGs' infiltration of as "*any case in which a natural person belonging to a criminal organization or acting on its behalf [...] invests financial and/or human resources to participate in the decision-making process of a legitimate business*" (Transcrime, 2017, p.19). The key feature of this definition is that a person tied to an OCG plays an active role in the decision process of a firm. Note, in particular, that the accrual of financial resources is *not* a necessary condition for infiltration. Given the definition of infiltration, the construction of the *Mappatura* involves two steps: (1) identification of individuals belonging to an OCG or acting on its behalf, (2) participation of such individuals in the decision-making process of legitimate businesses.

Identifying Individuals (Step 1) UIF is responsible for combating money laundering and terrorist financing. To help UIF perform its tasks, the law establishes disclosure requirements on financial intermediaries, supervisory authorities, administrative bodies, and professional associations. These entities transmit to UIF data on financial flows and other information (mainly) through suspicious transaction reports (STRs). UIF screens and analyzes STRs before transmitting them to investigative bodies. In 2022 alone, UIF received 155,426 STRs (UIF, 2022). STRs are the most comprehensive source of information available on transactions potentially linked to criminal activities.

All physical persons identified in STRs are searched for in several judicial and investigative databases on OCGs. This process produces a highly confidential list of around 40,000 matched individuals potentially implicated with OCGs. The most important source, accounting for around 90% of the matched individuals, are individuals of interest to the *Direzione Nazionale Anti-*

mafia (*DNA*). Established in 1991, the *DNA* coordinates *all* investigations relating to OCGs in Italy. The *DNA* list, therefore, includes *any* individual that Italian investigative bodies and judicial authorities consider of potential interest in investigating OCGs' activities. The remaining 10% are individuals arrested or investigated for involvement with OCGs in the World-Check database and others for whom UIF received information requests on OCG-related matters from judicial authorities, Italian investigative authorities, or foreign Financial Intelligence Units. In sum, the final list includes individuals who are under investigation for, or are reported in judicial documents connected to, OCG-related crimes *and* are also involved in suspicious financial transactions reported to UIF.

The comprehensiveness of this data source is crucial to gain an accurate picture of the phenomenon under consideration. First, belonging to an OCG is a crime on its own in Italy (article 416-bis of the penal code). However, in practice, it is difficult to prove in a court affiliation to a secret criminal organization. Many "Mafiosi" are therefore investigated, brought to court, and convicted for crimes other than 416-bis. Second, numerous investigations have highlighted how individuals who assist OCGs in infiltrating the legal economy –the so-called *zona grigia* (grey area) – are *not* members of the OCGs. These include not only figureheads, but also qualified professionals (e.g., accountants, lawyers, consultants) who knowingly support and act on behalf of OCGs, or independent entrepreneurs who knowingly do business (collude) with OCGs. The DNA list includes these individuals as well.

Identifying Firms (Step 2) Using unique social security identifiers, we match individuals on the list obtained in step 1 with the owners and directors of the universe of Italian firms extracted from the Infocamere database of the Italian Chamber of Commerce. A firm is then classified as infiltrated when it has at least one matched individual among its owners or directors. The firm's date of infiltration is the first year in which such a match occurs and can, of course, coincide with the year of creation of the firm.[8]

The *Mappatura* in Perspective The *Mappatura* identifies about 106,000 infiltrated firms, thereby casting a much wider net than previously possible. The current frontier in the field is the creative approach by Mirenda et al. (2022). Focusing on 'Ndrangheta infiltration in Central and Northern Italy, they de-

[8]Because of the extreme confidentiality of the underlying data, *none* of the authors had access to the list of individuals identified in Step 1. Step 2 was performed by separate staff at UIF. The identifier of the matched *firms* (but not of the matched *individuals*) was then shared with one of the UIF-affiliated authors of this paper. The other authors never had access to the data, including at the firm level.

fine infiltration as the presence of an owner or a director that carries the family name and birthplace typically associated with 'Ndrangheta families as reported in Dalla-Chiesa et al. (2014). Unlike our baseline definition, Mirenda et al. (2022) also includes firms with ownership ties to infiltrated firms. This approach yields a sample of about 9,000 infiltrated firms.

Calamunci and Drago (2020), Fabrizi and Parbonetti (2021), and Bianchi et al. (2022) study firms directly identified as infiltrated by OCGs according to investigative records – an alternative approach that arguably minimizes the likelihood of false positives. This approach, however, comes at the expense of a limited sample size of 450, 645, and 1,840 firms, respectively. Other useful benchmarks are the registry of confiscated firms, which consists of around 3,000 firms at the last stage of the confiscation process, and that of firms that are blacklisted for participation in public procurement, which consists of about 2,800 firms over the years 2016-2022. These two lists partially overlap and are not publicly available, to the best of our knowledge.[9]

Discussion of the *Mappatura* Given its comprehensiveness, inevitably the *Mappatura* might include some false positives. We take a conservative approach that alleviates such concerns. First, note that an individual must appear in STRs *and* in at least another source, such as the DNA database, to make it to our list. We thus exclude individuals in STRs even when signaled as potentially connected to OCGs unless at least another data source backs up such suspicion. For the same reason, we also exclude firms flagged in STRs as potentially related to OCGs but without a matched individual on their board. Second, the *DNA* list assigns a risk indicator articulated in five levels. Based on extensive conversations with relevant practitioners, the *Mappatura* omits individuals with the lowest risk (score equal to 1), who may include simple acquaintances and others informed of the facts being investigated. Our results are robust to excluding individuals with a score equal to 2 and with a missing score.[10]

On the other hand, the *Mappatura* certainly suffers from false negatives. By definition, it misses individuals that do not generate any STRs. *DNA* data are extremely sensitive and UIF is only informed whether individuals reported in STRs are in the database. For confidentiality reasons, we also do not know how

[9] Decarolis et al. (2024) match a confidential dataset from AISI (Italy's domestic intelligence and security agency) that identifies individuals suspected of various crimes to firm-level records, without the ability to separate OCG involvement from other crimes.

[10] In principle, another potential source of false positives is individuals under investigation who end up 'clean'. While accurate data are not available, practitioners consider this to be a minor issue given the precautionary steps described above.

many individuals are on the *DNA* list. Furthermore, our methodology, which identifies infiltrated firms through the presence of owners or administrators, naturally misses infiltration cases in which none of these roles in the firm is involved. For example, a firm under the grip of an OCG through usury or extortion will not be classified as infiltrated, unless an individual tied to the OCG appears on the firm's board.

3 How Pervasive is OCGs' Infiltration?

The Incidence of Infiltration The *Mappatura* identifies 106,122 infiltrated firms over our sample period (2005-2020). This number corresponds to $\approx 2\%$ of all corporations and partnerships in Italy (for confidentiality reasons, we must exclude sole proprietorships). Table 1 describes the differences between infiltrated firms and other firms in the economy. Column 1 (2) reports the average characteristics of firms in the *Mappatura* (other firms). Infiltrated firms are younger, larger, and more likely to be corporations. These patterns hold both unconditionally and conditional on province-sector fixed effects (column (3)).

Since firms in the *Mappatura* are larger than the typical Italian firm, we weigh infiltration by firm size. Social security records (INPS) reveal that firms potentially connected to OCGs account for 8 to 10% of private sector employment in Italy, excluding sole proprietorships (which, however, represented only 9% of private sector employment in 2019). Similarly, these firms account for 10 to 14% of the revenues of firms in CERVED – a population that includes all (non-financial) firms required to disclose balance sheets and financial statements and that accounted for 58% of aggregate value added in Italy in 2019. It is worth emphasizing that our definition of infiltration does *not* imply that firms in the *Mappatura* are *controlled* by OCGs. It does imply, however, that a person allegedly connected to OCGs has a prominent role in the firm.

Geographic Distribution Figure 1 Panel A maps the incidence of *Mappatura* firms across Italian provinces. Unsurprisingly, there is a higher incidence in the home regions of the main OCGs in the South of Italy. The highest incidence is found in the provinces of Reggio Calabria, Vibo Valentia, Crotone, and Catanzaro in Calabria – home of the *'ndrangheta*. Substantial infiltration is also recorded in Napoli, Caserta, and Salerno in Campania and across most provinces in Sicily. Nevertheless, around two-thirds of infiltrated firms are found in the more prosperous regions in the Center-North of Italy. Table A1, Panel A in the Appendix explores the correlation between institutional factors

Table 1: Infiltrated Firms vs Non-Infiltrated Firms

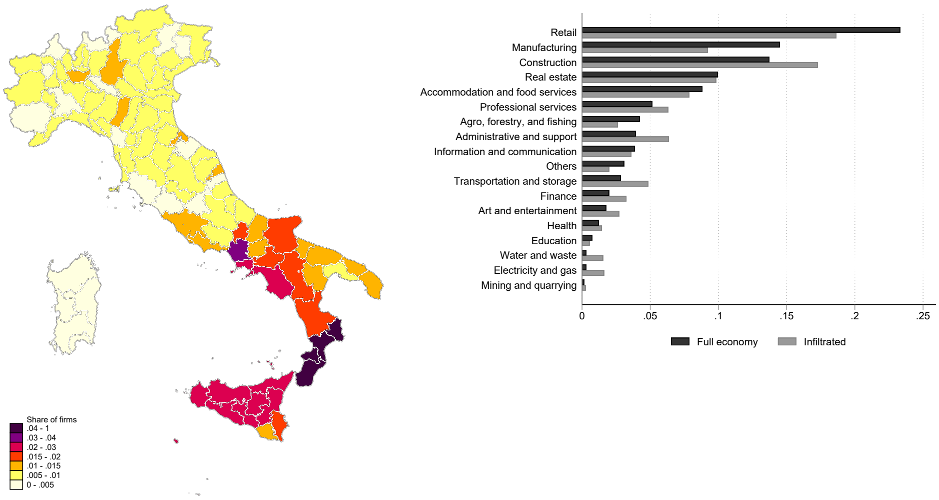
	Mappatura, all	Full economy	*Mappatura* dummy coefficient	*Mappatura*, born clean	*Mappatura*, born infiltrated	Born-clean dummy coefficient
	(1)	(2)	(3)	(4)	(5)	(6)
All Firms in *Infocamere*						
Year of birth	2002.7	2000.6	1.23***	1997.1	2008.2	-10.62***
No. employees	16.3	2.9	13.68***	27.8	5.4	18.53***
=1 if corporation	0.717	0.517	0.15***	0.749	0.686	0.04***
=1 if partnership	0.175	0.407	-0.17***	0.137	0.211	-0.05***
No. directors	2.1	1.6	0.52***	2.4	1.9	0.46***
No. owners	3.0	2.5	0.52***	3.6	2.5	0.97***
No. auditors	4.7	4.3	0.50***	5.0	4.1	0.82***
Number of firms	106,122	5,224,062		51,935	54,187	
Firms in *Cerved*						
Assets (IHS)	6.9	6.0	0.97***	7.3	6.3	0.86***
Revenues (IHS)	6.7	5.9	0.82***	7.0	6.3	0.55***
Payroll (IHS)	5.3	4.8	0.68***	5.7	4.9	0.57***
Number of firms	64,388	2,079,674		33,231	31,157	

Notes: The unit of observation is a firm for all statistics on the table. Column (1): average characteristics of *Mappatura* firms. Column (2): average characteristics of all firms in the economy (Infocamere). Column (3): estimated difference between *Mappatura* firms and other firms, conditional on province-by-sector fixed effects, for each row variable. Columns (4) and (5) report average characteristics for *born-infiltrated* and *born-clean* firms respectively. Column (6): estimated difference among *Mappatura* firms between born-clean firms and born-infiltrated firms, conditional on province-by-sector fixed effects, for each row variable. The Infocamere sample corresponds to the universe of Italian firms, excluding sole proprietorship. Assets, revenues, and payroll (inverse hyperbolic sine transformation) are only available for firms reporting balance sheet information (Cerved database). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

and the presence of infiltrated firms across Italian provinces. The share of infiltrated firms is higher in provinces with lower income per capita (consistent with evidence, cited above, that OCGs originated in less developed regions and subsequently stifled economic development), lower degrees of social capital (measured by blood donation and trust), slower courts (both consistent with hypothesis in the literature, see, e.g., Gambetta, 1996), and a higher share of the population with family names from the regions of origins of Italian OCGs (a proxy that presumably correlates with OCGs ability to control the territory). Interestingly, when we include all variables at once (Column 7), the correlation with income per capita becomes positive. This is consistent with the idea that, once we control for the ”supply” of OCGs (through proxies for social capital and territory of origins), income per capita captures the ”demand” – i.e., business opportunities – for OCGs.

Figure 1: Geographic and Sectorial Distribution of Infiltration

(a) Geographic distribution **(b)** Distribution by sector



Notes: These figures present the geographic and sector-level distribution of firms in the *Mappatura*. Panel A presents the share of infiltrated firms relative to the number of firms in the province. Panel B presents the share of firms in each sector for the whole economy (black bars) and for firms in the *Mappatura* (grey bars).

Sectoral Distribution Figure 1, Panel B presents the share of firms in each sector for the whole economy (black bars) and for firms in the *Mappatura* (grey bars). Unlike the stark geographic divide, OCGs' infiltration is quite balanced across sectors. We confirm the previous literature view that certain sectors, particularly those that deal with the public administration, are particularly vulnerable to infiltration: there are disproportionately high shares of infiltrated firms in construction, transportation and storage, and utilities, with waste collection, treatment, and disposal the sector with the highest incidence of infiltrated firms (11.5%). At the same time, the Figure illustrates how the wide net of the *Mappatura* recovers a distribution of infiltration across sectors that is broadly representative of the nationwide economic structure. This suggests that while traditional sectors in which OCGs can deploy their criminal expertise are certainly relevant, they are far from being the sole, and perhaps the main, destination of OCGs' infiltration into the legal economy. This observation motivates our conceptual framework in the next Section.

4 Conceptual Framework

This Section turns to the motives of infiltration. These motives are unobservable and must be inferred from the data. We introduce a new taxonomy of motives and a conceptual framework that guides our empirical analysis.

4.1 Motives of Infiltration

The illegal activities of OCGs – i.e., their core business – generate large amounts of liquidity. Several considerations suggest that such liquidity is not exclusively reinvested in illegal activities. For example, the risk of confiscation tampers returns from illegal activities: a basic risk-diversification argument suggests allocating some of the revenues into safer – from the point of view of the threat of enforcement – assets. Furthermore, OCGs cannot advertise their illegal products (which limits the customer base) and must also rely on few, trusted, counterparts in the illegal economy: (the threat of) violence is not a perfect substitute to formal contract enforcement since it attracts unwelcome attention from the police. Finally, enjoying the fruits of the illegal business also requires legal means of payment other than cash.

In sum, the optimal portfolio management of a large OCG requires investing at least part of the profits generated by illegal activities in legal assets. Infiltration of legitimate firms is one of the asset classes available to OCGs. Although estimates are difficult to come by, our calculations from official records suggest that firms account for only 10-20% of the value of all assets confiscated to OCGs in Italy, most of them being real estate. But this may be just a reflection of the fact that while the value of real estate may be insensitive to confiscation, that of a business may collapse following confiscation. Furthermore, it might be difficult to confiscate infiltrated firms, particularly if infiltration occurs through board members rather than ownership of equity shares.

Why would OCGs infiltrate legitimate firms? Answering this question is critical to designing effective investigative and law enforcement strategies; indeed, the extensive criminal and policy literature has sought answers. Several taxonomies have been developed from detailed analysis of samples of infiltrated firms identified in investigations (see, e.g., Arlacchi, 2010, Commissione Antimafia, 2018, De Simoni, 2022). For example, Parbonetti (2021) distinguishes between *supporting* firms (empty shells used to mask illegal activities), *cartiere* firms ("paper mills" that specialize in false invoicing), and *star* firms that display superior economic performance thanks to their connection with OCGs.

These classifications provide a useful taxonomy of infiltration motives from the perspective of investigative bodies and are consonant with the legal definition of the infiltrated firm (see Fn. 2). Both the analysts' taxonomy and the criminal code definition share the idea that somehow OCGs involve the firm in their criminal activities. For example, in what we label the *competitive* motive, they leverage criminal expertise to benefit the firm. Alternatively, in what we label the *functional* motive, they use the firm to support criminal activities. We call legitimate firms infiltrated with these motives ”contaminated”, meaning that the OCG contaminates the legal activity of the infiltrated firm with criminal activity or criminal methods.

However, involving the firm in criminal activities increases the risk of detection. It is thus conceivable that OCGs may infiltrate legitimate firms purely to generate safer, albeit smaller, pecuniary returns or other non-pecuniary benefits that cannot be obtained through criminal activities or by investing in other legal assets. For example, OCGs can obtain relatively high and safe returns if they can access "private equity" like investment opportunities. Furthermore, being involved in the management of a firm, even (or, perhaps, especially) if it is not ”contaminated” by criminal activities, opens the door to connections to other firms, public administration, politicians, etc., fostering the relational capital of the OCG. These connections can be useful to identify further opportunities – both legal (like for any other entrepreneur), and illegal. We label this new and previously underappreciated bundle of motives ”pure” – with the understanding that there is nothing pure about it, since ultimately infiltration happens by, or on behalf of, OCGs. In our terminology, ”pure” simply means that the legitimate firm itself is not "contaminated" with criminal activities.[11]

We aim to infer the underlying motives of infiltration from the economic behavior of a large sample of infiltrated firms identified in the *Mappatura*. To do so, we put forward a parsimonious conceptual framework encompassing the ”contaminated” and the ”pure” motives. The infiltration of OCGs into legitimate businesses is a complex phenomenon and, inevitably, our framework entails a degree of simplification. At the same time, the approach maps testable predictions to different motives of infiltration thereby enabling quantification of their relative prevalence in the *Mappatura*, overcoming the limitations of studies based on smaller samples of firms identified in judicial investigations.

[11] Money laundering, therefore, belongs to the *contaminated*, rather than the *pure* motive. See Section 6 for a discussion of money laundering and how it relates to the *pure* motive.

4.2 Set-Up

Benchmark As a benchmark, consider an entrepreneur with access to an investment opportunity that returns output $y = \theta f(k)$. θ is the entrepreneur's talent, $f(k)$ an increasing and strictly concave function, k a bundle of inputs, including capital, labor, and materials, that must be financed and that we refer to as the firm's scale of operation. The entrepreneur has no funds and borrows from a competitive lending market (banks) at an interest rate of $1 + r$. The entrepreneur solves

$$\max_{k} \quad \Pi(k) = \theta f(k) - (1 + r)k, \tag{1}$$

yielding a unique solution, k^*, implicitly defined by $\theta f'(k^*) = (1 + r)$. Denote with Π^* the profits at k^*. Critically, k, y, and Π are observable in the data.

The Infiltrated Firm The key feature of infiltration is that a person tied to an OCG plays an active role in the firm's decision process. We focus on whether to "contaminate" the firm with the OCG's criminal activities. On the one hand, doing so benefits the OCG in potentially many ways. On the other hand, it increases the risk that the firm is confiscated. If there were no benefits unless the firm is involved in criminal activities, all infiltration would be of the *contaminated* type. To capture the idea that OCGs may want to infiltrate the firm without involving it in criminal activities – the *pure* motive – we assume that the OCG has unlimited funds that yield a pecuniary return $(1 + i) < (1 + r)$. This introduces a pecuniary motive for infiltration.[12] We focus on how infiltration changes the firm's *demand* for bank finance and assume that banks do not adjust their *supply* of funds to the firm in response. The infiltrated firm thus borrows k_b from banks at interest rate $(1 + r)$ and obtains $k_m \geq 0$ (given our definition, either equity or debt) from the OCG. Like before, denote the firm's scale of operation as $k = k_m + k_b$. The OCG solves

$$\begin{aligned}
\max_{k,k_m,\mathbf{I}^c,\mathbf{I}^f} \quad & V^m(k, k_m, \mathbf{I}^c, \mathbf{I}^f) = \\
& = \underbrace{((1 + \lambda \mathbf{I}^c)\theta f(k) - (1 + r)k + (r - i)k_m)}_{\Pi^m}(1 - \rho(\mathbf{I}, k_m, k^*)) + \mathbf{I}^f \gamma C(k), \\
s.t. \quad & k_m \leq k; \mathbf{I}^c, \mathbf{I}^f \in \{0, 1\}, \mathbf{I}^c + \mathbf{I}^f \leq 1.
\end{aligned} \tag{2}$$

Comparing (1) and (2) reveals how the infiltrated firm differs from a legitimate

[12]Recall, however, that the accrual of financial resources is *not* a necessary condition for infiltration according to our definition. For simplicity, we abstract from non-pecuniary benefits in the model and discuss them in detail in Section 6.

firm facing similar investment opportunities and markets. First, the indicator functions $\mathbf{I}^c$ and $\mathbf{I}^f$ capture whether the firm is contaminated with the OCG's criminal activities or not. $\mathbf{I}^c$ captures a *competitive* motive in which the OCG's criminal activities enhance the firm's performance, e.g., a firm that acquires larger market shares by threatening competitors or wins public procurement contracts by corrupting public officials. All else equal, we expect λ to be higher in the OCGs' home regions, where they exert firmer control over local institutions and society. $\mathbf{I}^f$, instead, captures a *functional* motive in which the firm's scale of operation k is distorted away from profit maximization to pursue criminal activities that yield payoff $\gamma C(k)$, with $C'(k) > 0$ a natural assumption. Firms that produce false invoicing to facilitate money laundering or hire extra workers to acquire consensus in, and control over, a certain territory are examples of this typology. Again, λ will likely be higher in the OCGs' home regions. The *contaminated* motive emerges when the solution to the program in (2) yields $\mathbf{I} = \max\{\mathbf{I}^c, \mathbf{I}^f\} = 1$.[13] In contrast, the *pure* motive emerges when the solution entails $\mathbf{I} = 0$ and the firm is kept separate from the OCG's criminal activities.

Second, infiltration introduces a risk of confiscation, $\rho(\mathbf{I}, k_m, k^*) \in (0, 1)$.

Assumptions:

1 $\partial\rho(1, k_m, k^*)/\partial z \geq 0$ *for* $z = k_m, k^*$;

2(a) $\rho(1, k_m, k^*) > \rho(0, k_m, k^*)$ *for all* $k_m, k^* \geq 0$,

2(b) $\partial\rho(1, k_m, k^*)/\partial z > \partial\rho(0, k_m, k^*)/\partial z$ *for all* $z \geq 0, z = k_m, k^*$

Assumption 1 states that the risk of confiscation is increasing in k^* and k_m: the larger the firm and the larger the OCG's involvement in the firm, the more likely that the firm ends up under the investigative radar.[14] Assumption 2(a) states that, for all levels of k^* and k_m, the risk of confiscation is higher in the *contaminated* motives than in the *pure* motive. Assumption 2(b) states that the scale of the firm and the OCG's involvement increases the likelihood of detection more when the firm is a *contaminated* than when it is *pure*. These assumptions appear natural: for example, an OCG is more likely to attract attention when it threatens a competitor, or it wins a public contract rigging a procurement auction, relative to when it simply provides cheaper finance to an entrepreneur.

[13]For simplicity, we assume that the firm is involved in either of the two motives, but not both contemporaneously, i.e., $\mathbf{I}^c + \mathbf{I}^f \leq 1$.

[14]The assumption that $\rho(\cdot, k^*)$ depends on the undistorted scale of the firm, k^*, rather than k, captures the intuition that larger firms are under more scrutiny – e.g., because they must disclose more information – without overly complicating the algebra.

Finally, the OCG's funds, k_m, lowers the cost of capital of the firm by $(r - i)k_m$. This happens in both the *contaminated* and *pure* motive. The OCG potentially invests $k_m > 0$ only if $r > i$. If that was not the case, the *pure* motive would never arise in the model, since k_m increases the risk of detection.

4.3 Observable Behaviour of Different Infiltration Motives

We first describe how different infiltration motives alter the firm's scale of operation, k, and sources of finance, k_b/k, relative to each other and the benchmark. We then solve for the optimal motive, $\mathbf{I}^c$ and $\mathbf{I}^f$, as a function of θ.[15]

Operational Scale In an interior solution, the first-order condition w.r.t. k is

$$(1 + \lambda \mathbf{I}^c)\theta f'(k) + \frac{\mathbf{I}^f \gamma C'(k)}{(1 - \rho(\mathbf{I}, k_m, k^*))} = (1 + r). \tag{3}$$

Relative to non-infiltrated firms, the firm's operating scale k, expands in the *contaminated* infiltration ($\mathbf{I} = 1$) but not in the *pure* motive ($\mathbf{I} = 0$). *Competitive* infiltration, $\mathbf{I}^c = 1$, increases the returns from investing in the firm, while *functional* infiltration, $\mathbf{I}^f = 1$, distorts the firm's operation scale to support criminal activities. In contrast, k is undistorted under *pure* infiltration.

Implication 1 *Contaminated infiltration increases firm's scale of operation k and revenues y, pure infiltration does not.*[16]

Sources of Finance The two motives of infiltration also differ in the sources of finance employed by the firm. The first-order condition w.r.t. k_m,

$$\Pi^m \frac{\rho'(\mathbf{I}, k_m, k^*)}{(1 - \rho(\mathbf{I}, k_m, k^*))} = r - i, \tag{4}$$

highlights the key trade-off: funds from the OCG, k_m, yield a marginal benefit $r - i$; but increase the risk of detection. As noted above, in the *pure* motive, the scale of the firm k does not change – this motive is thus characterized by a *substitution* of bank finance k_b with OCG funds. *Pure* infiltration takes advantage of the lower cost of capital supplied by the OCG to substitute more expensive sources of finance for the firm. If the OCG's funds were not cheaper, there wouldn't be a reason to invest.

[15] We focus on an interior solution ($k_b > 0$) as the likely most relevant empirical case. Even with multiple sources of finance, the OCG might substitute those that bring more scrutiny to the firm (e.g., bank loans) rather than the most expensive ones.

[16] The *competitive* and *functional* motives can also be distinguished: the former always increases profits Π^m, the latter potentially decreases them.

The *contaminated* infiltration has more nuanced implications. OCG's funds increase the risk of detection more than in the *pure* motive (Assumption 2), and thus – all else equal – *contaminated* infiltration implies a lower k_m relative to *pure*. In fact, if $\rho'(1, 0, k^*)$ is sufficiently large, the *contaminated* firm exclusively relies on external sources of finance ($k_m = 0$), as the benefit of cheaper finance, $(r-i)$, does not compensate for the increased risk of detection.

Implication 2 *Pure infiltration substitutes external sources of finance with internal ones, contaminated infiltration less so, if at all.*[17]

4.4 The Choice of Infiltration Motives

When does the OCG prefer one motive over the other? That is, for which parameters does the solution (2) involve $\mathbf{I} = 1$ as opposed to $\mathbf{I} = 0$? Since motives are unobservable, an answer to this question is necessary to derive testable predictions. While a full characterization of the solution lies beyond the scope of our analysis, the trade-off we have highlighted implies a clear comparative static for the optimal choice as a function of θ.

Figure 2 illustrates the OCG's preferred motive of investment as a function of θ, focusing on an interior solution in which all three motives – *functional*, *competition* and *pure* – arise.[18] All else equal, the *contaminated* motives have a higher risk of detection relative to the *pure* (Assumption 2). The *contaminated* motives are thus more likely when the value of the confiscated firm is not too large, i.e., for low values of θ. Between the two *contaminated* motives, however, the benefits of the *competitive* motive $(1 + \lambda)$ are complementary to θ, while those of the *functional* motive (captured by γ) are not. Hence the *functional* motive is chosen for firms with lower θ relative to the *competitive* one. The firm's size, i.e., its scale of operation, k, monotonically increases in θ. The model thus characterizes the infiltration motive as a function of the firm's size.

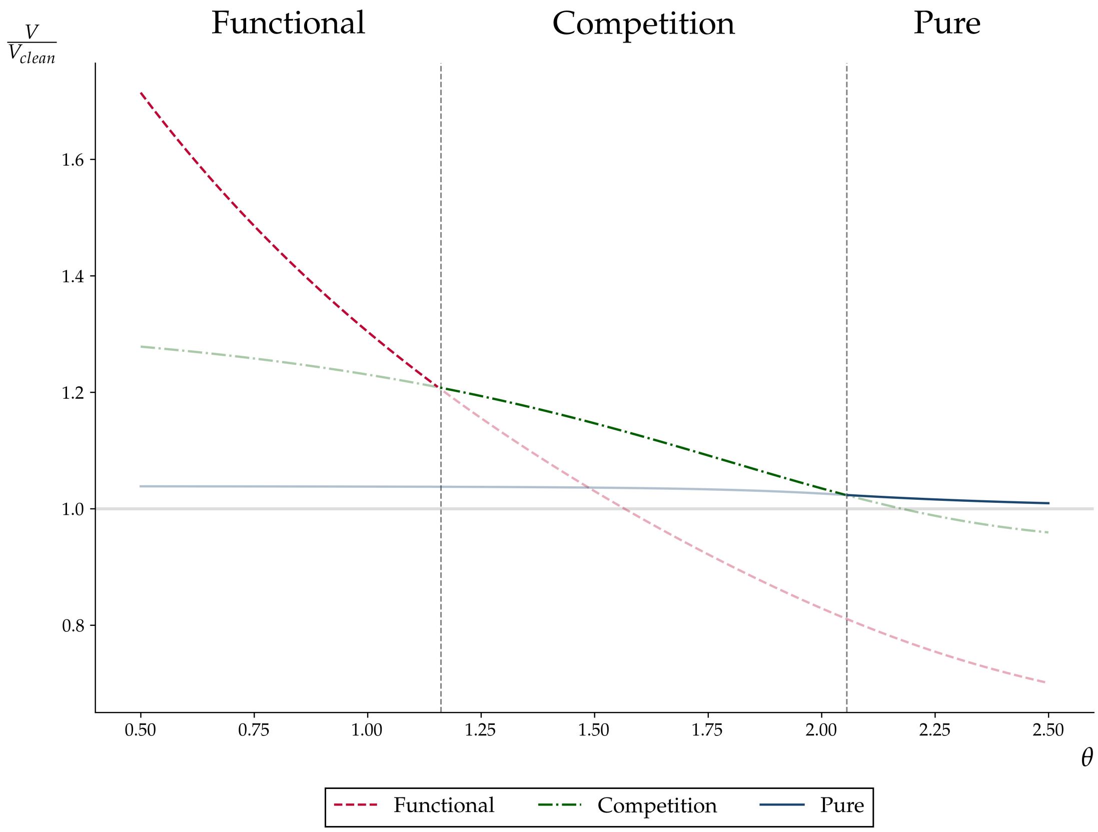
Implication 3 *The functional motive is chosen for small firms, the competitive motive for medium-sized ones, the pure motive for the largest firms.*

Combining implications 1 and 2 (which describe observable hallmarks as-

[17] Consistently with this, Parbonetti (2021) note that *cartiere* – firms used for false invoices – often accumulate large debts with the tax authority before shutting down; while the superior performance of *star* firms often attracts *more* bank finance.

[18] The Figure also focuses on a case in which infiltration always yields a higher payoff than the profits earned by a clean entrepreneur with identical θ and r (i.e., $V^m > \Pi^*$). This, of course, need not be the case, e.g., if $i \to r$, or if $\rho(\cdot, 0, k^*)$ is sufficiently large, infiltration can yield a lower payoff at least over some ranges of θ.

Figure 2: Firm's Size and Infiltration Motives (Comparative Statics)



Notes: The Figure reports the value of the solution of (2) relative to the value of (1) for different values of θ. The values resulting from the optimal choice of k and k_m are plotted for $\mathbf{I}^c = 1, \mathbf{I}^f = 1$, and $\mathbf{I}^c + \mathbf{I}^f = 0$. The solution to program (2) is given by the upper envelope of the three curves. Functional forms: $f(k) = \frac{k^{\epsilon}}{\epsilon}$; $C(k) = \frac{k^{\varsigma}}{\varsigma}$; $\rho(I, k_m, k^*) = (1 - I)\rho_1(k_m) + I\rho_2(k_m)\rho_2(k^*)$, with ρ_i a logistic function with supremum L_i, growth rate g_i, and midpoint ξ_i. Parameters: $\lambda = \frac{1}{2}$; $r = 0.15$; $i = 0.025$; $\epsilon = \frac{1}{4}$; $\gamma = \frac{1}{3}$; $\varsigma = \frac{1}{3}$ $L_1 = \frac{1}{10}$; $g_1 = 2$; $\xi_1 = e$; $L_2 = \frac{3}{\sqrt{10}}$; $g_2 = \frac{5}{4}$; $\xi_2 = 0$.

sociated with each motive) with implication 3 (which characterizes for which firms each motive arises) yields a rich set of testable predictions. Figure A1 illustrates the main testable predictions by reporting observable firms' outcomes – the risk of confiscation $\rho(\cdot)$, the scale of operation k, and the sources of finance k_m/k (panel (c)) – for different values of θ using the parameters in Figure 2. Furthermore, inspection of (2) reveals that – ceteris paribus – the *contaminated* motives are more likely to emerge the higher λ and γ, i.e., in OCGs' home regions. Similarly, infiltration – particularly of the *pure* type – is more likely to occur ($V^m > \Pi^*$) for firms that struggle to borrow from banks (higher r). These comparative statics thus yield additional testable predictions.

5 Motives of Infiltration: Evidence

We leverage the *Mappatura* to test the model's predictions and infer the relative prevalence of different infiltration motives. Section 5.1 introduces a distinction between *born-infiltrated* and *born-clean* firms. At the time of infiltration, *born-infiltrated* firms are smaller than *born-clean* firms. Implication 3 suggests that *born-infiltrated* firms are more likely to reflect the riskier *contaminated* motives relative to *born-clean* firms. We document patterns in geographic diffusion and risk of confiscation consistent with this hypothesis. We then explore Implications 1 and 2 separately on the two groups of firms, since they differ in the empirical strategies available to construct a suitable comparison group of non-infiltrated firms. Section 5.2 examines *born-clean* firms, Section 5.3 *born-infiltrated* ones. The *Mappatura* supports *all* the testable implications of the model, with the hallmarks of the *contaminated* motives detected on *born-infiltrated* and smaller *born-clean* firms, and the hallmarks of the *pure* motive on larger, *born-clean* firms.

5.1 Prediction 1: *Born-Infiltrated* versus *Born-Clean* firms

The data reveal an important distinction: 51% of all firms in the *Mappatura* are *born-infiltrated* (i.e., the presence of the individual tied to the OCG is detected when the firm is established) and 49% are *born-clean* (i.e., the individual connected to an OCG enters in the firm at a later date). Our framework is not micro-funded to distinguish born-clean and born-infiltrated firms: conditional on parameters, the two modes of infiltration yield the same solution.[19] Table 1, however, shows that *born-infiltrated* firms are significantly smaller than *born-clean*: they have fewer employees, assets, and revenues. Combined with Implication 3, these differences in size imply:

Prediction 1 *Relative to born-clean firms, born-infiltrated firms are more likely to reflect a contaminated motive than a pure motive. Born-infiltrated firms thus (a) are at a higher risk of confiscation (higher $\rho(\mathbf{I}, k_m, k^*)$), and (b) are more prevalent in OCGs' home regions (higher γ and λ)* .

Table 2 confirms prediction 1(a): *born-infiltrated* firms have a higher likelihood of being confiscated. We obtain data on all firms confiscated to OCGs by

[19] To see why, consider a born-clean firm set-up by an entrepreneur. Upon infiltration, the OCG maximizes $V^m(k, k_m) - T$, s.t. the entrepreneur's participation constraint, $T \geq \Pi^*$. This yields the same solution of the born-infiltrated firm in (2).

Table 2: Born-Infiltrated vs. Born-Clean: Confiscation Risk

	(1)	(2) All firms	(3)	(4)	(5) Infiltrated Firms Only	(6)
Born-Infiltrated	0.013***	0.012***	0.012***	0.003**	0.006***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Born-Clean	0.013***	0.009***	0.009***			
	(0.001)	(0.001)	(0.001)			
Observations	5,175,704	1,169,379	1,169,379	80,586	28,340	28,340
R-squared	0.036	0.040	0.040	0.191	0.195	0.195
YOB × Province × Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	No	No	Yes
Infiltrated at birth firms	50689	22356	22356	43254	18430	18430
Infiltrated after	48770	12015	12015	37332	9910	9910
Mean dep var	0.0007	0.0007	0.0007	0.0162	0.0133	0.0133
p-value diff. in coefficients	0.998	0.0221	0.0228			

Notes: This table presents the difference in risk of confiscation across groups of firms. Columns 1 and 4 use the universe of firms from infocamere, while other columns focus on the sample of firms in the CERVED dataset. All regressions include year of birth by industry (2-digit) by province of birth fixed effects. Controls are total assets, revenue, and number of employees, all measured at birth. Robust standard errors clustered at province of birth-year of birth level are presented in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

judiciary authorities in Italy. Firms in the *Mappatura* are around 30 times more likely to have been confiscated relative to other firms (columns (1) through (3)). Among firms in *Mappatura*, *born-infiltrated* firms are twice as likely to have been confiscated than *born-clean* firms (columns (4) through (6)).

Appendix Figure A2 confirms prediction 1(b). The *contaminated* motives are more likely when benefits from the OCG's involvement are large: when λ or γ are high. This is presumably so where OCGs' have a higher degree of socio-economic control, i.e., in their regions of origin. The Figure reports the share of firms that are *born infiltrated* in the population of infiltrated firms. While we observe significant variation across all of Italy, born-infiltrated firms are clearly over-represented in the regions of origin in the South of Italy (Appendix Table A2 confirms this in a regression framework). Appendix Table A1, Panel B also finds that the share of born-infiltrated firms is higher in provinces with a lower institutional development (lower economic activity, lower trust, slower courts, higher prevalence of family names from the regions of origins of Italian OCGs).

These patterns give us confidence that the distinction between *born-clean* and *born-infiltrated* firms captures different underlying motives. We now verify, separately for the two samples, the hallmarks of the corresponding motives described in Implications 1 and 2.

5.2 Prediction 2: Infiltration Motives in *Born-clean* Firms

Born-clean firms are relatively large. The model therefore implies that these firms predominantly reflect the *pure* motive. Implications 1 and 2 yield:

Prediction 2(a)-(b) *Born-clean firms mainly reflect the pure motive. Infiltration is associated (a) with no significant change in the firm's sales y and scale of operation k, and (b) with substitution of the sources of finance away from bank loans (lower k_b/k).*

5.2.1 Empirical Approach

Born-clean firms are, by definition, observed both before and after infiltration. To test the prediction, we thus compare infiltrated firms' outcomes around the date of infiltration relative to a suitably constructed comparison group of non-infiltrated firms within a difference-in-differences framework. A firm's *date of infiltration* is defined as the year in which an OCG-linked individual first joined the firm either as an owner or director. As such, an empirical issue arises by which, *by definition*, infiltration of born-clean firms coincides with changes in the firm's ownership and management, which are likely to arise during special circumstances of a firm's life and could be associated with changes in the firm's performance and operations.

We propose an empirical approach that compares infiltrated firms to non-infiltrated firms that *also* experience an inflow of a new owner or director. That is, we first define an *inflow event* as the year in which a new owner or director joins the firm. In the case of several such occurrences during our sample period, we denote as the focal *inflow event* the one in which the greatest number of new owners or directors joined the firm, selecting the earliest year in case of ties. We then explicitly account for such *inflow events* within the DID framework.

Of course, our approach is not intended to identify the "causal" effect of infiltration on the firm. Infiltration is certainly not a random event – by definition, it involves an individual with links to OCGs, while other *inflow events* do not. In our setup, the DID simply offers a convenient way to describe the data. In particular, pre-trends observed *before* infiltration occurs are also potentially informative about the strategy pursued by OCGs. For example, our model implies that infiltration might target firms that struggle to borrow from banks. The model thus speaks both to pre-trends themselves and to dynamics in post-infiltration outcomes. Interpreted through the model's lens, and bearing these caveats in mind, the DID approach is intuitive and leads to pretty clear results regarding OCGs' motives of infiltration.

Empirical Specifications We compare the evolution of firm outcomes following infiltration vis-a-vis the evolution following a non-criminal inflow estimating the following regression:

$$y_{ipst} = \alpha_i + \alpha_{st} + \alpha_{pt} + \beta_1 \times \text{Post I}_{it} + \beta_2 \times \text{Post INF}_{it} + \epsilon_{ipst}, \tag{5}$$

where i, p, s, and t stand for firm, province, sector, and year. The dummy variable Post I_{it} takes value one after firm i has experienced the inflow event, regardless of whether it involved individuals tied to OCGs or not. The dummy variable Post INF_{it}, instead, takes value one after firm i is infiltrated. Our key parameter of interest, β_2, captures the differential change in outcome y after an infiltration compared to the differential change for a non-criminal inflow. That is, β_1 captures the change in outcomes following a non-criminal inflow event while the equivalent effect for an infiltration event is given by $\beta_1 + \beta_2$.

We include three sets of fixed effects: α_i are firm-level fixed effects that absorb time-invariant heterogeneity across firms, α_{st} are sector-year fixed effects that absorb any sector-level (39 2-digits sectors) heterogeneity that changes over time, and α_{pt} are province-year fixed effects that capture any province-level (107 provinces) time-varying heterogeneity. Finally, ϵ_{ipst} is an error term arbitrarily correlated over time within a firm.

We also investigate firm outcomes' dynamics around non-criminal inflows and infiltration by estimating an event study specification:

$$y_{ipst} = \alpha_i^k + \alpha_{st}^k + \alpha_{pt}^k + \sum_{j \in \{-5,\ldots,-2,0,\ldots,5\}} \gamma_j^k \cdot D_{i,t-j}^k + \epsilon_{ipst}^k, \tag{6}$$

for $k \in \{I, INF\}$, where I stands for a non-criminal inflow event and INF stands for infiltration. $D_{i,t-j}^I$ is a dummy variable equal to one if firm i experienced a non-criminal inflow event $t-j$ periods ago. $D_{i,t-j}^{INF}$ is a dummy variable equal to one if firm i experienced an infiltration event $t-j$ periods ago. We estimate the dynamic effects γ_j^I and γ_j^{INF} separately, in samples that include firms that experienced the relevant type of inflow and firms that never experienced any inflow as controls. Our coefficients of interest, the difference $\gamma_j^{INF} - \gamma_j^I$, describe changes in outcomes around infiltration relative to a clean inflow event. Section 5.2.3 describes several robustness checks to specifications (5) and (6).[20]

[20]Table A3 shows that regular inflows and infiltrations differ in firms' characteristics in the year before the event. These differences, however, are relatively small when considering standardized average differences.

5.2.2 Empirical Results

Prediction 2(a): Operational Outcomes Figure 3 illustrates the dynamics of firm operational outcomes around infiltration and other inflows. The left panels report estimates for γ_j^{INF} and γ_j^I from specification (6), while the right panels report the difference between the two, $\gamma_j^{INF} - \gamma_j^I$. Table 3, Panel A, reports estimates from the static specification in (5).

The left panels in Figure 3 appear to suggest that infiltration raises the firm's scale of operation. However, dynamics around a non-criminal inflow event are very similar to those around infiltration. The left panels, then, show that after accounting for the inflow event itself, changes associated with infiltration are indistinguishable from zero. These patterns emerge for several operational outcomes: revenues, employment, payroll, and intermediate inputs. In line with the evidence from Figure 3, Table 3, Panel A, reveals large changes in operational outcomes associated with inflows (estimated β_1 range in $0.11 - 0.23$), while changes associated with infiltration (β_2) are small and close to zero.

The results indicate that infiltration of *born-clean* firms is not associated with changes in the firm's operation relative to any firm that experiences an inflow of new owners or managers. This is potentially consistent with the *pure* motive which, indeed, the model suggests should be particularly prevalent among these firms. To confirm this prediction, however, we need to consider the financial position of the firm, since the *pure* motive implies changes in these outcomes. Unlike the operational outcomes, we are going to find significant changes in the firm's financial position.

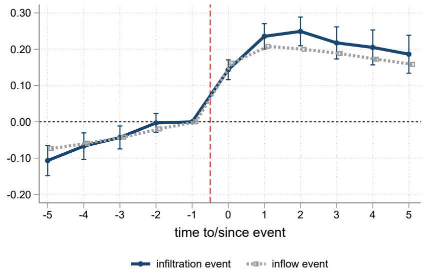
Prediction 2(b): Sources of Finance Figure 4 illustrates the dynamics of the firm's financial position around infiltration and clean inflows. As before, the left panels report estimates for γ_j^{INF} and γ_j^I from equation (6), while the right panels report the difference between the two. Panel B of Table 3 reports the corresponding estimates from the static specification in (5).

Consistent with prediction 2(b), the Figure and the Table uncover a stark difference. Detailed data from the credit registry at the Bank of Italy reveal that infiltration is associated with a substitution away from bank loans relative to regular inflow events: normal inflows coincide with a significant increase in bank borrowing, and infiltration does the opposite. These patterns emerge both on the extensive (panels (a) and (b)) and intensive (panels (c) and (d)) margin.

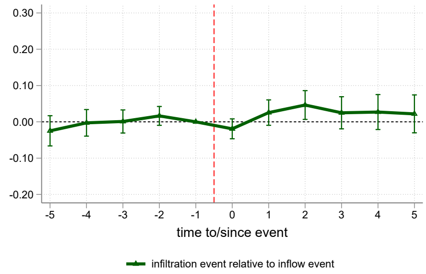
The decline in bank loans upon infiltration deserves a more careful discussion. In principle, the reduction in bank lending to infiltrated firms could stem

Figure 3: Infiltration, Revenues, and Operational Outcomes

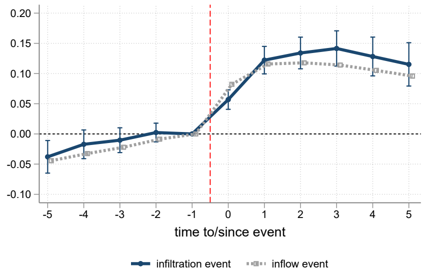
(a) Revenue: infiltration vs. inflow



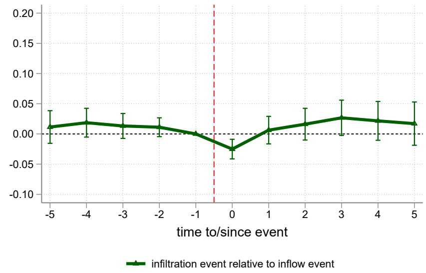
(b) Revenue: Δ infiltration and inflow



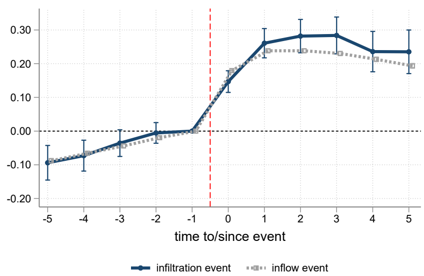
(c) Number of employees: infiltration vs. inflow



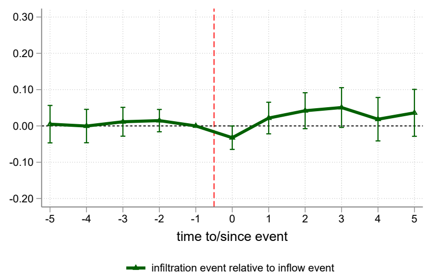
(d) Number of employees: Δ infiltration and inflow



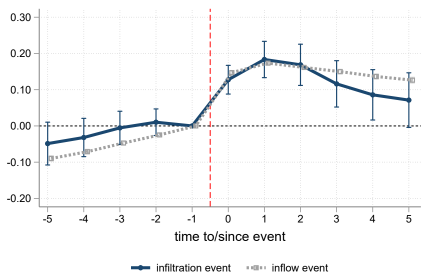
(e) Payroll: infiltration vs. inflow



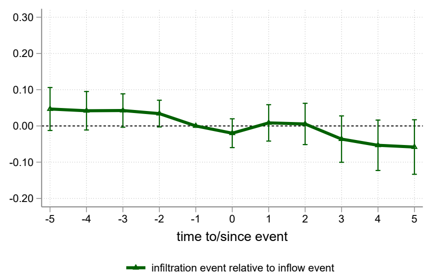
(f) Payroll: Δ infiltration and inflow



(g) Intermediate inputs: infiltration vs. inflow



(h) Intermediate inputs: Δ infiltration and inflow



Notes: Left panels: Point estimates and 95% confidence intervals for parameters γ_j^{INF} and γ_j^{NC} from equation (6). Right panel: Difference between γ_j^{INF} and γ_j^{NC} estimates. For all firms with an inflow event, we include observations from -5 and +5 years around the event. The specification includes sector-year and province-year fixed effects. All outcome variables are in the inverse hyperbolic sine form except for the first row, which is a dummy variable.

Table 3: Infiltration of Born-Clean Firms

	(1)	(2)	(3)	(4)
Panel A: Operational outcomes				
Dep. variable	Revenues	No. Employees	Payroll	Inputs
Post Infiltration	0.014	-0.006	0.006	-0.041*
	(0.016)	(0.010)	(0.020)	(0.022)
Post Any Inflow	0.203***	0.112***	0.234***	0.180***
	(0.002)	(0.001)	(0.002)	(0.002)
No. observations	9,758,931	9,758,931	9,758,931	9,758,931
No. firms	1,555,154	1,555,154	1,555,154	1,555,154
No. infiltrated firms	17,708	17,708	17,708	17,708
No. inflow event firms	828,022	828,022	828,022	828,022
Panel B: Financial outcomes				
Dep. variable	=1 any bank loans	Bank loans if >0	Receivables	Cash
Post Infiltration	-0.040***	-0.206***	0.112***	-0.010
	(0.004)	(0.059)	(0.015)	(0.018)
Post Any Inflow	0.028***	0.151***	0.162***	0.070***
	(0.001)	(0.007)	(0.002)	(0.002)
No. observations	9,758,931	5,217,909	9,758,931	9,758,931
No. firms	1,555,154	829,942	1,555,154	1,555,154
No. infiltrated firms	17,708	10,231	17,708	17,708
No. inflow event firms	828,022	492,759	828,022	828,022

Notes: This table presents the point estimates from equation (5). The sample excludes firms born infiltrated. For all firms with an inflow event, we include observations from -5 and +5 years around the event. Post Infiltration$_{it}$ takes the value one after a firm i is infiltrated, while Post Any Inflow$_{it}$ takes the value one after firm i is infiltrated or experiences a non-criminal inflow event. All columns include sector-year and province-year fixed effects. All outcome variables are in inverse hyperbolic sine form except for the dummy variable `=1 any bank loans`. Standard errors clustered at the firm level. * p<0.1, ** p<0.05, *** p<0.01.

from a *demand* or from a *supply* channel. The *demand* channel is the one emphasized by our conceptual framework: the firm no longer needs to borrow funds from the bank due to the cheaper sources of finance brought in by the OCG.[21] On the *supply* side, however, banks' response is *a priori* ambiguous. On the one hand, banks might reduce lending to a firm they perceive to have been infiltrated or involved in dodgy deals. Our model also suggests that OCGs might target firms that struggle to borrow, in which case the reduced borrowing from banks might continue a pre-trend in which the supply of funds to the firm is drying up. On the other hand, it is worth noting that the bank might become more willing to lend to a firm whose financial position has improved due to the entry of new sources of funds.

The *supply* channel is unlikely to account for the entire reduction in bank

[21] In practice, the *demand* channel might mask a further motive. Through suspicious transactions reports (STRs), banks are the backbone of the financial crime enforcement system. Infiltrated firms might thus prefer to shy away from interactions with banks to limit scrutiny, rather than saving on the costs of capital. This argument is particularly plausible among larger firms and justifies our focus on the interior solution of the model.

borrowing. Panel (e) and (f) in Figure 4 and column (3), panel B of Table 3, reveal that infiltrated firms increase their commercial credit following infiltration—i.e., they become net suppliers of working capital for other firms in their supply chain. While the supply of funds from banks might have partially dried up, the overall sources of finance available to the firm have expanded.

Consequently, the liquidity position of the firm improves. The involvement of owners with large amounts of cash, or of administrators with links to potential financiers, should be associated with an increase in the liquidity of the firm. The last row in Figure 4 supports this prediction, as illustrated by the trend reversal and jump in cash holdings around the time of infiltration. Interestingly, the static difference-in-difference specification in Column 8 of Table 3 misses this effect (it is indeed negative, but not statistically different from zero). This happens because infiltrated firms display a negative pre-trend relative to the control group of firms who experience a non-criminal inflow, consistent with OCGs targeting, or being accepted by, firms that are experiencing (potentially temporary) liquidity problems. Combined with the earlier results on operational outcomes, these findings indicate that OCGs target firms likely in *financial* but not in *economic* distress.

Predictions 2(c): Heterogeneity Among Born-Clean Firms Taken together, the results so far support the predictions of the model: *born-clean* firms present the hallmarks of the *pure* motive. To the extent that *born-clean* firms include a mix of motives, the higher risk of detection associated with the *contaminated* motive is relatively less costly for small and young firms. The dynamics of infiltration for these firms, therefore, should present the hallmarks of the *contaminated* motive, rather than the *pure* one. That is, the model implies:

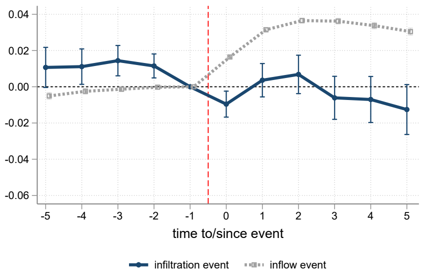
Prediction 2(c) *Within born-clean firms, smaller firms reflect a contaminated motive: the firm's revenues y and scale of operation k increase.*

Table 4 supports this prediction. While non-criminal inflows coincide with larger changes in operational outcomes on smaller firms, these heterogeneous estimates are significantly larger for infiltrated firms. Similar results are found for firms infiltrated when younger (unreported).[22]

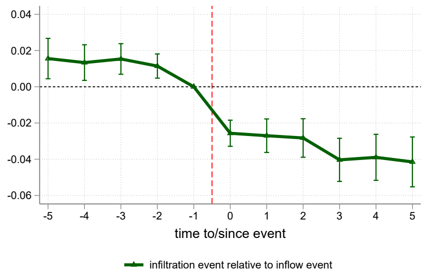
[22]Columns (1) and (4) find a reduction in revenues and inputs after infiltration on larger firms. A simple extension of the model in which infiltrated firms face an uninsurable confiscation risk when choosing k also matches this evidence. For simplicity, however, we have assumed a risk-neutral OCG.

Figure 4: Infiltration and Financial Position

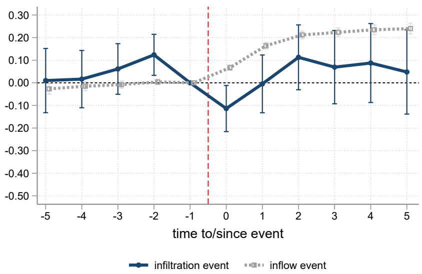
(a) Bank loans, ext. margin: inf. vs. inflow



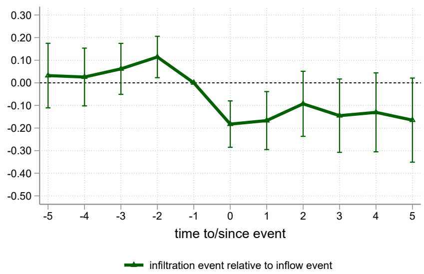
(b) Bank loans, ext. margin: Δ inf. and inflow



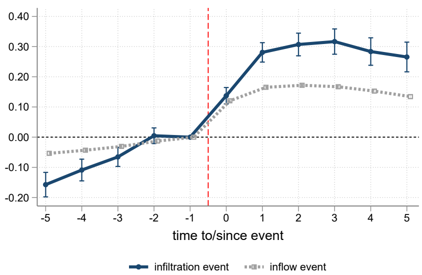
(c) Bank loans, int. margin: inf. vs. inflow



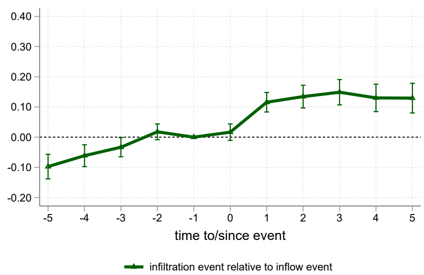
(d) Bank loans, int. margin: Δ inf. and inflow



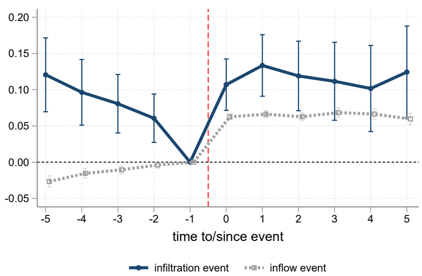
(e) Receivables: inf. vs. inflow



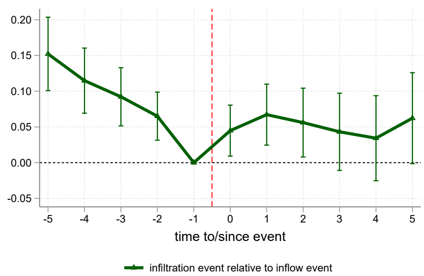
(f) Receivables: Δ inf. and inflow



(g) Cash holdings: inf. vs. inflow



(h) Cash holdings: Δ inf. and inflow



Notes: Left panels: Point estimates and 95% confidence intervals for parameters γ_j^{INF} and γ_j^{NC} from equation (6). Right panel: Difference between γ_j^{INF} and γ_j^{NC} estimates. For all firms with an inflow event, we include observations from -5 and +5 years around the event. The specification includes sector-year and province-year fixed effects. All outcome variables are in the inverse hyperbolic sine form except for the first row, which is a dummy variable.

Table 4: Born-Clean Firms: Heterogeneity

Dep. variable	Revenue (1)	No. Employees (2)	Payroll (3)	Inputs (4)
Post Infiltration × Small	0.208***	0.007	0.045	0.187***
	(0.031)	(0.022)	(0.039)	(0.045)
Post Infiltration	-0.060***	0.018	0.035	-0.086**
	(0.023)	(0.017)	(0.028)	(0.035)
Post Any Inflow × Small	0.273***	0.157***	0.313***	0.366***
	(0.003)	(0.003)	(0.005)	(0.006)
Post Any Inflow	-0.006**	-0.008***	-0.007*	-0.101***
	(0.003)	(0.002)	(0.004)	(0.005)
Observations	9,758,931	9,758,931	9,758,931	9,758,931
Mean dep variable	6.371	1.435	3.617	4.184
Number of infiltrated	17708	17708	17708	17708
Number of inflow firms	828022	828022	828022	828022

Notes: This table presents the point estimates from equation (5). The sample excludes firms born infiltrated. For all firms with an inflow event, we include observations from -5 and +5 years around the event. Post Infiltration$_{it}$ takes the value one after a firm i is infiltrated, while Post Any Inflow$_{it}$ takes the value one after firm i is infiltrated or experiences a non-criminal inflow event. *Small* takes value one for firms with assets less than 2m Euro one year before the inflow event. All columns include sector-year and province-year fixed effects. All outcome variables are in inverse hyperbolic sine form except for the dummy variable `=1 any bank loans`. Standard errors clustered at the firm level. * p<0.1, ** p<0.05, *** p<0.01.

5.2.3 Robustness and Comparison with Mirenda et al. (2022)

Appendix B explores the robustness of our findings along several dimensions. First, under heterogeneous treatment effects, the two-way fixed-effects (TWFE) model suffers from "bad" comparisons if later treated units are used as a control for early treated units. In our context, the number of never-treated units (i.e., firms that did not experience either an infiltration or an inflow event) is large, thus reducing this concern. Indeed, the share of estimates with negative weights (i.e., coming from such "bad" comparisons as in De Chaisemartin and d'Haultfoeuille (2020)) is small (0% for infiltration and 6% for other inflow events). Nevertheless, we estimate both a stacked-panel model that compares infiltrated firms to those that experience an inflow in the same year, as in Cengiz et al. (2019) (Figure B1 and Table B1), as well as the static model that retains the staggered adoption dimension, as in Wooldridge (2021) (Table B2). Results across both the operational (sales, employees, payroll, materials) and the financial (loans from banks, commercial credit, liquidity) outcomes are robust to both specifications. Second, we confirm robustness to several changes in our definition of infiltration (Tables B3 and B4) and to alternative sample restrictions (Table B5). Results on both operational and financial outcomes are robust to extending the *Mappatura* to include firms owned by infiltrated firms, or to (further) restricting the *Mappatura* to firms identified through matched individuals with a high risk in the *DNA* list, or to excluding firms in the South, or to keeping the sample constant across outcomes.

Our analysis of infiltration of *born-clean* firms parallels Mirenda et al. (2022). Since we find radically different results from theirs, it is important to understand where the differences arise. In principle, the difference could arise from *i*) a different empirical proxy for infiltration (i.e., the use of the *Mappatura* rather than family surnames associated with OCGs), or *ii*) a different sample (our sample includes all of Italy while Mirenda et al. (2022) focus on firms in the North); *iii*) our comparison to other firms that also experience an inflow event. The latter drives the difference. Focusing on revenues (the main outcome shared by the two analyses), Appendix Figure A3 and Table A4 show that introducing our correction in Mirenda et al. (2022)'s infiltration proxy and sample recovers our result. A further difference with Mirenda et al. (2022) is that, while they restrict their analysis to born-clean firms to implement DID specifications, we now turn our attention to born-infiltrated firms, for which our model suggests we might find evidence consistent with a different infiltration motive.

5.3 Prediction 3: Infiltration Motives in *Born-Infiltrated* Firms

Studying *born-infiltrated* firms is important for two reasons. First, these firms account for about half of all firms in the *Mappatura* and for 7.2% of all assets invested in newly created firms in the typical year. Second, and crucially, the model suggests that the motive of infiltration for these firms might be different than for *born-clean* firms. Born-infiltrated firms are relatively small. Implications 1, 2, and 3 imply:

Prediction 3(a)-(b) *Born-infiltrated firms predominantly reflect a contaminated motive and are thus associated with 3(a) a higher scale of operation at birth, 3(b) lower performance and worse selection (lower θ) in the case of functional ($\mathbf{I}^f = 1$) relative to the competition motive ($\mathbf{I}^f = 1$).*

While prediction 3(a) immediately follows from implications 1, 2, and 3, prediction 3(b) requires an explanation. Within the *contaminated* motive, the *competitive* motive is associated with a better firm performance, while the *functional* motive with a decrease since, holding constant θ, k is distorted away from its profit-maximizing level as the OCG also takes into account the criminal payoff $\gamma C(k)$. This suggests that relative to a non-contaminated firm in the same market, the infiltrated firms of the *functional* type survive even with low productivity θ, provided it is sufficiently well-selected on γ. While not formally modeled, this selection argument naturally emerges in an extension that models entry and survival along the lines of Melitz (2003).

Empirical Approach Unlike born-clean firms, born-infiltrated firms are, by definition, not observed before infiltration. The difference-in-differences framework is thus not feasible. To test the predictions, we compare *born infiltrated* firms to firms established in the same year, province, and sector. This comparison necessarily bundles how infiltration alters the firm's operation with the process of entrepreneurial selection. To investigate selection, we borrow the empirical specification from Banerjee and Munshi (2004) study of capital misallocation in India. In their framework, a group of "insiders" has better access to capital than "outsider" entrepreneurs. Insider firms do not need to be as good as the financially constrained "outsiders" to survive and are thus negatively selected. While initially larger, these firms display a lower performance over time. The analogy with infiltrated firms of the *functional* type is perfectly fitting: as they are selected for their criminal payoff (γ), they do not need to have a θ as high as clean firms to survive in the market. Borrowing from Banerjee and Munshi (2004), we estimate:

$$y_{it} = \beta_1 \times BI_i + \beta_2 \times BI_i \times Age_{it} + \alpha_{psb} + \alpha_t + \epsilon_{itpsb} \tag{7}$$

where i, t, p, s, b stand for firm, year, province, sector, and cohort. BI_i is a dummy that takes value equal one if firm i was born-infiltrated, while Age_{it} is the age of firm i in year t. α_{psb} are year of birth by province of birth by industry fixed effects, while α_t are year fixed effects. Age_{it} is collinear with cohort and year effects and is thus absorbed in the specification.

Empirical Results Table 5 reports the results and validates the model's predictions. Across the same operational outcomes explored above – revenues, employment, payroll, and inputs – panel A reveals that *born-infiltrated* firms are initially larger ($\beta_1 > 0$) but subsequently grow at a slower pace than firms born in the same year-province-sector ($\beta_2 < 0$). These findings align with the negative selection implied by the model. Furthermore, total assets increase over time, while profitability decreases (Table A5). For completeness, Panel B considers financial outcomes. Recall that the model doesn't make clear-cut predictions on the sources of finance for the *contaminated* motive. Born-infiltrated firms have significantly higher levels of liquidity and *more* loans than firms born in the same year-province-sector. This last result starkly contrasts with the findings for *born-clean* firms and further highlights the different motives of infiltration across the two groups of firms.

Table 5: Born-Infiltrated Firms at Birth and Over Time

	(1)	(2)	(3)	(4)
Panel A: Operational outcomes				
Dep. variable:	Revenue	No. Employees	Payroll	Inputs
Born infiltrated	0.165***	0.156***	0.199***	0.228***
	(0.022)	(0.009)	(0.018)	(0.020)
Born infiltrated × Age	-0.014**	-0.008***	-0.024***	-0.031***
	(0.005)	(0.002)	(0.004)	(0.005)
Panel B: Financial outcomes				
Dep. variable:	=1 any bank loans	Bank loans if >0	Receivables	Cash
Born infiltrated	0.020***	0.704***	0.627***	0.257***
	(0.003)	(0.028)	(0.018)	(0.014)
Born infiltrated × Age	-0.001	0.020**	0.034***	-0.005
	(0.001)	(0.005)	(0.004)	(0.004)
Observations	6,126,878	6,126,878	6,126,878	6,126,878
Infiltrated firms	22455	22455	22455	22455
Mean dep var (Panel A)	4.848	0.985	2.629	3.215
Mean dep var (Panel B)	0.338	4.274	4.243	3.123
YOB × Province × Industry FE	Yes	Yes	Yes	Yes

Notes: Born-Infiltrated is a dummy taking value = 1 if the firm was born infiltrated. *Age* measures the firm's age every year. The sample includes all firms in the CERVED dataset. All regressions include year of birth by province of birth by 2-digit industry fixed effects, and year fixed effects. All outcome variables are in inverse hyperbolic sine form, except for column 1 Panel B which is a dummy. For consistency, we include all firms in the CERVED dataset. Imposing a positive revenue restriction yields an even stronger negative selection. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1.

6 Discussion and Policy Implications

The evidence supports our distinction between the *contaminated* and *pure* motives of infiltration. This distinction departs from the dominant idea in the literature that infiltration is always contaminated with criminal activities. Such characterization applies well to smaller and medium-sized firms, often directly established by the OCGs. Many firms in the *Mappatura*, however, are already large and well-established when infiltration occurs. The behavior of these firms is in line with the implications of what we have labeled the *pure* motive, in which the infiltrated firm remains disconnected from criminal activities. *Born-clean* firms, which are more likely to reflect this *pure* motive, account for 85% of the assets of firms in the *Mappatura*. This previously undetected *pure* motive might thus be a significant, if not the predominant, motive of infiltration. We conclude by discussing non-pecuniary benefits likely associated with the *pure* motive, the relationship between the *pure* motive and money laundering, and the policy implications of our findings.

6.1 The Non-Pecuniary Benefits of Infiltration

For simplicity, the model assumes that the benefit of *pure* infiltration is pecuniary: OCGs earn higher risk-adjusted financial returns by investing liquidity in legitimate firms rather than reinvesting in criminal businesses or hoarding funds ($r > i$). This raises the question of why direct involvement in the management or ownership of Italian firms is necessary to achieve these returns, compared to alternative assets or legitimate businesses in safer jurisdictions. While OCGs may access "private equity"-like investments with high returns in Italy, most infiltration —especially among *born-clean* firms that likely reflect the pure motive — occurs through administrators, not owners. These considerations (alongside the arguments in footnotes 15 and 21) suggest that higher financial returns may not be the sole benefit of *pure* infiltration.

We conjecture that, alongside the financial returns, *pure* infiltration is likely motivated by non-pecuniary benefits as well. Crucially, these private benefits are distinct from those in the *functional* motive, in which the OCG benefits from the crimes the firm is directly involved with. In contrast, in the *pure* motive, the OCG seeks private benefits that (*i*) are disconnected from criminal activity, and (*ii*) can only be acquired by being involved in the operation of large firms.

The acquisition of what we label "relational capital" – a web of relationships with important actors in the legal economy (such as board members of other large firms, industry associations, public administration, politicians, consultants, etc.) – is likely a key non-pecuniary benefit of infiltrating large, well-established, firms. Such relationships are valuable to OCGs in many ways (e.g., providing information about small and medium-sized firms needing liquidity, or public policies that present lucrative opportunities for corruption or fraud) but can *only* – or almost exclusively – be accessed through direct involvement in large, legitimate, businesses. Other investments do not provide these benefits.

This hypothesis naturally leads to testable implications. While the benefits are unobservable, relational capital can be proxied by the extent and composition of connections in the network of board members, implying differences in such connections for both infiltrat*ed* firms and infiltrat*ing* individuals.

Connections of Infiltrat*ed* Firms The more connected a firm's board members are to other firms, the greater the value of the relational capital acquired through infiltration. Similarly, connections to politicians might yield particularly valuable relational capital. If so, infiltrat*ed* firms – and particularly so born-clean ones – should differ in the extent and type of their connections:

Prediction 4(a)-(b) *Infiltration – particularly of large born-clean firms – targets firms with board members that are (a) more connected to other firms, (b) more likely to be politicians.*

Table A6 in the Appendix shows that OCGs infiltrate firms with board members more connected to other firms. Using the individual-level panel of owners and administrators for born-clean infiltrated firms and inflow-event firms, we compute the average number of connections per board member in the year before the infiltration or the inflow event. Relative to other firms in the same province, sector, and year, board members of infiltrated firms have more connections to other firms (columns (1) to (3)) and particularly so to other infiltrated firms (columns (4) to (6)). On average, board members of infiltrated firms have three times as many connections as board members of firms experiencing a clean inflow. Similarly, the average board member has 0.2 connections with infiltrated firms, whereas infiltrated firms' corresponding figure is around 6.[23]

Table 6 shows that infiltration is more likely to target firms that have politicians on their boards. We match owners and administrators of firms with the universe of previously elected politicians to Italian municipal and regional councils, the Italian national parliament, and the European parliament. We focus on born-clean firms, for which we can identify the presence of elected politicians on the firm's board in the year before the infiltration occurs. As with the analysis of born-clean firms in Section 5.2, we compare firms that experience infiltration relative to firms that experience a non-criminal inflow.[24]

Infiltration targets firms more likely to have a politician on their boards, compared to non-criminal inflows (column (1)). This pattern holds at any political level: municipal (column (3)), regional (column (5)), national parliament

[23]Permutation tests with randomly selected placebo lists of OCG-linked individuals and corresponding placebo *Mappatura* firms in Figure A4 reveal that these patterns are neither an artifact of infiltrated firms being larger nor the result of chance.

[24]Unreported results, show that born-infiltrated firms are also more likely to have a politician on their board when established, relative to firms born in the same year, sector, and province. However, for confidentiality reasons, we do not know whether the politician on the board is him/herself on the list of connected individuals. This pattern is thus harder to interpret. We can instead observe born-clean firms the year before infiltration occurs and be confident that the politician is not him/herself connected to OCGs.

Table 6: Infiltration and political connections

Dep. variable:	Political connection											
	Any		Local politician		Regional politician		Italian Parliament		EU Parliament		EU Lobby	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Infiltrated	2.284***	5.552***	2.146***	5.282***	1.347***	2.472***	0.134***	0.407***	0.057***	0.176**	0.002	0.011
	(0.203)	(0.577)	(0.200)	(0.569)	(0.127)	(0.407)	(0.028)	(0.103)	(0.019)	(0.071)	(0.005)	(0.020)
Infiltrated × Small		-4.892***		-4.703***		-1.761***		-0.355***		-0.156**		-0.012
		(0.611)		(0.602)		(0.423)		(0.105)		(0.072)		(0.020)
Observations	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420	1,026,420
Mean dep variable	9.167	9.167	8.940	8.940	2.858	2.858	0.0415	0.0415	0.0237	0.0237	0.00205	0.00205
Number of infiltrated	22818	22818	22818	22818	22818	22818	22818	22818	22818	22818	22818	22818

Notes: This table presents the relationship between infiltration and political connections. We compare infiltrated firms to firms that experience a non-criminal inflow-event, focusing on the year before the inflow. The dependent variable takes the value of one hundred if at least one of the owners, administrators, or auditors of the firm was a (previously) elected politician. Columns 1 and 2 present the correlation between infiltration and the prevalence of any political connections, columns 3 and 4 focus on connections to local politicians, 5 and 6 to regional politicians, 7 and 8 to members of the Italian Parliament, 9 and 10 to members of the EU Parliament. Controls include total revenue, total assets, liquidity, industry, province, and year of event fixed effects. *Small* takes value one for firms with assets less than 2m Euro in the year before the event. Columns 11 and 12 focus on whether the firm registered lobbying activities at the EU parliament. Robust standard errors are presented in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

(column (7)), and European parliament (column (9)). Columns (2), (4), (6), (8), and (10) show that, at all political levels, the correlation between infiltration and the presence of a politician on the board is stronger for larger firms, defined as in Table 4. This is consistent with the hypothesis that the acquisition of relational capital is particularly important in the *pure* infiltration, which we have argued is more likely among large born-clean firms. For these firms, the estimated coefficients imply a stronger correlation between infiltration and political connection as we move up the political hierarchy: the estimated size effect is larger for members of parliament than for regional and local politicians.[25]

Connections of Infiltrat*ing* Individuals The *pure* motive needs individuals that are competent, look clean and respectable, and yet are tied to, and can be trusted by, the OCG. These individuals are in scarce supply and therefore must be leveraged across several, and larger, infiltrations – as in the large, born-clean, firms on which we have detected the *pure* motive.[26] That is:

Prediction 4(c) *Infiltrating individuals themselves have more connections to other firms, especially infiltrated ones.*

Table A7 shows that *within* boards, individuals connected to OCGs sit on

[25]Columns (11) and (12) explore whether the firms engaged in lobbying activities at the EU Parliament. Although the estimates are noisy – only 0.2% of firms have engaged in lobbying – large born-clean firms are about 5 times more likely to have engaged in lobbying activities at the EU level than equally sized firms that also experienced inflow events.

[26]Conversely, the *contaminated* motives hinge on criminal activities, for which there are economies of scale: once a person has committed a crime (e.g., threatened a competitor through violence), (s)he might just as well do so across other transactions. This implies that the *contaminated* motive appears in small and *born-infiltrated* firms. Contaminating the firm with criminal activities, instead, is more costly for a "clean" entrepreneur. This provides a micro-foundation for the distinction between the infiltration of born-clean and born-infiltrated firms.

more boards than other board members of the same infiltrated firm. The Table reports estimates from individual-level regressions that control for firm-times-year fixed effects. Relative to other board members in the same firm, individuals connected to OCGs sit on the boards of more firms (columns (1)–(3)) and on more boards of firms with other individuals connected to OCGs, both on the intensive (columns (4)–(6)) and extensive (columns (7)–(9)) margins. Results are robust regardless of whether individuals are weighted equally or by the number of board-years they appear in.

In sum, the *pure* motive may reflect OCGs' desire and ability to engage with key players in the legal economy, such as large enterprises, politically connected individuals, public administrators, and high-profile service providers (e.g., lawyers, accountants, consultants). This hypothesis was corroborated by prosecutors and investigators particularly familiar with recent trends regarding the 'Ndrangheta – the OCG that likely accounts for most firms in the *Mappatura*.

6.2 The Relationship of Pure Infiltration to Money Laundering

Money laundering (ML) converts the proceeds of crime into assets with a legitimate appearance, enabling their indefinite retention or use in further criminal activities.[27] ML typically involves three stages (Gilmore, 2004): *placement* (introducing funds into the financial system), *layering* (disguising their origin), and *integration* (deployment of funds that appear legitimate).

The criminology literature, primarily based on investigations and court cases, often finds evidence of rudimentary ML schemes, mostly limited to the *placement* and *layering* stages (Riccardi and Reuter, 2024), with little evidence of *integration*. This contrasts with reports by the Financial Action Task Force (FATF) and journalistic investigations (*Offshore Leaks*, *Panama Papers*), which highlight sophisticated ML techniques.

The *pure* infiltration discussed in this paper is conceptually distinct from ML. Common ML techniques involving legitimate firms often rely on false invoicing, where revenues are artificially inflated, leading to increased tax liabilities and eventual bankruptcy, or offset by collusive input purchases – in which case, the colluding supplier then features an anomalous increase in revenues. The most complex schemes may thus travel along the supply chain up to small firms subject to minimal accounting scrutiny, or exploit VAT fraud in

[27] See, e.g., the definition in the UK's *The Proceeds of Crime Act, 2002 (POCA)*. Legal definitions vary; for instance, Article 648-bis of the Italian penal code criminalizes the replacement, transfer, or transaction of money or goods derived from crime to obscure their origin.

cross-border transactions. In contrast, under *pure* infiltration, the firm itself is not being used to launder money; indeed, we have found no evidence of such anomalous behavior among larger, born-clean firms

Nevertheless, the *pure* motive of infiltration identified in this paper contributes to our understanding of ML by identifying cases in which legitimate firms are potentially used for the final, and most elusive, *integration* stage. For example, a legitimate-seeming entrepreneur tied to OCGs might invest criminally derived funds that have already been *placed* and *layered*. If such a firm is not directly involved in laundering through mechanisms like false invoicing, we label the infiltration as *pure*, despite potential links to prior ML stages that could be uncovered by a careful investigation.

6.3 Policy Implications and conclusions

The distinction between *contaminated* and *pure* motives carries critical policy implications at both operational and strategic levels in combating OCGs. First, if most infiltration links directly to criminal activities, focusing scarce investigative resources on counter-crime measures can detect infiltration and curtail OCGs' returns. However, as our findings suggest, a substantial portion of infiltration – including legitimate-seeming human capital and laundered funds – often operates independently of underlying crimes. These undetected flows expand OCG investment opportunities and connections, posing significant risks. Unlike *contaminated* motives, which have higher detection risks due to direct links to illegal activities, the *pure* motive relies on covert financial operations and relational capital investments, reducing exposure. Consequently, our results highlight a potential misallocation of resources: investigative efforts might have to prioritize analyzing financial transactions and relationships rather than exclusively targeting overt criminal activities. Addressing *pure* motives requires enhancing financial analysis expertise and identifying at-risk professionals colluding with OCGs.

Second, and within the anti-money laundering apparatus, the design of monitoring systems, leniency programs, and screening algorithms, depends on the extent to which OCGs involve legal firms in criminal activities or not. On the monitoring front, the evidence calls for a significant upgrade of the collaboration provided by specific categories of reporting agents – such as auditing firms and consultants – who are typically closer, by the nature of their function, to the firm's economic and financial developments and changes in governance.

An important avenue for future work is to understand the role of these individuals in facilitating *pure* infiltration. A key distinction between the *contaminated* and the *pure* infiltration is that in the former there likely is a victim (e.g., the competitor who was threatened or who lost the public procurement contract because of corruption), in the latter, there isn't (by definition, the entrepreneur is willing to accept the 'pact with the devil' and benefit from the OCG's cheaper finance and relational capital). Leniency programs for financial crimes connected to OCGs might thus have to be strengthened, with appropriate incentives, to fight infiltration of the *pure* motive. Furthermore, our evidence provides insights that are relevant for the design and optimization of algorithms used to detect infiltrated firms (see, e.g., Cariello et al., 2024 for an operational contribution). These algorithms are increasingly used by public investigative agencies as well as private entities (e.g., banks) to monitor transactions, detect suspicious operations, and be compliant with anti-money laundering regulations.

Finally, a more concerning implication emerges when considering a potentially important source of benefits from *pure* infiltration – political connections. Our results suggest that the *pure* infiltration might significantly increase the economic power of OCGs, as they present themselves with a totally clean and faultless image. They can thus interact freely and develop connections with the main economic players (managers of large enterprises, high-profile consultants, public officers making decisions on tenders, and politicians). This accumulation of "relational capital" can have far-reaching consequences. Given the well-known influence of economic lobbies on the legislative process in modern democracies (see, e.g., Bertrand et al., 2014, Bertrand et al., 2023), this economic power can become, over time, political power: i.e., OCGs can ultimately affect the law-making process (e.g., the design of anti-money laundering and financial regulation) thus strengthening and perpetuating their grip on the economy and society. Our findings align with alarms raised in recent years by the Italian intelligence and security agencies (see, e.g., DIS, 2019) and resonate with analyses of the Latin America's case. As the *Financial Times* recently put it, "*while mafias don't seek to overthrow the government, they seed "parallel powers" – networks of corrupt politicians, judicial officials, and bureaucrats – that disable the state's law enforcement capacity*" (FT, 2024).

References

Acemoglu, D., G. De Feo, and G. D. De Luca (2020). Weak states: Causes and consequences of the Sicilian Mafia. *The Review of Economic Studies 87*(2), 537–581.

Alesina, A., S. Piccolo, and P. Pinotti (2019). Organized crime, violence, and politics. *The Review of Economic Studies 86*(2), 457–499.

Arlacchi, P. (2010). *La mafia imprenditrice.* Il saggiatore.

Backhaus, J. (1979). Defending organized crime? *The Journal of Legal Studies*, 623–631.

Bandiera, O. (2003). Land reform, the market for protection, and the origins of the Sicilian mafia: Theory and evidence. *JLEO 19*(1), 218–244.

Banerjee, A. and K. Munshi (2004). How efficiently is capital allocated? Evidence from the knitted garment industry in Tirupur. *The Review of Economic Studies 71*(1), 19–42.

Bertrand, M., M. Bombardini, R. Fisman, F. Trebbi, and E. Yegen (2023). Investing in influence: Investors, portfolio firms, and political giving.

Bertrand, M., M. Bombardini, and F. Trebbi (2014). Is it whom you know or what you know? An empirical assessment of the lobbying process. *American Economic Review 104*(12), 3885–3920.

Bianchi, P. A., A. Marra, D. Masciandaro, and N. Pecchiari (2022). Organized crime and firms' financial statements: Evidence from criminal investigations in Italy. *The Accounting Review 97*(3), 77–106.

Blattman, C., G. Duncan, B. Lessing, and S. Tobón (2024). Gang rule: Understanding and countering criminal governance. *Review of Economic Studies*.

Bofondi, M., L. Carpinelli, and E. Sette (2018). Credit supply during a sovereign debt crisis. *Journal of the European Economic Association 16*(3), 696–729.

Brown, Z., E. Montero, C. Schmidt-Padilla, and M. Sviatschi (2024). Market structure and extortion: Evidence from 50,000 extortion payments. *Review of Economic Studies*.

Buchanan, J. M. (1973). A defense of organized crime. *The economics of crime and punishment 119*, 119.

Buonanno, P., R. Durante, G. Prarolo, and P. Vanin (2015). Poor institutions, rich mines: Resource curse in the origins of the Sicilian mafia. *The Economic Journal 125*(586), F175–F202.

Calamunci, F. and F. Drago (2020). The economic impact of organized crime infiltration in the legal economy: Evidence from the judicial administration of organized crime firms. *Italian Economic Journal 6*, 275–297.

Cariello, P., M. De Simoni, and S. Iezzi (2024). A machine learning approach for the detection of firms linked to organised crime in Italy. *Quaderni dell'antiriciclaggio UIF 22*.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics 134*(3), 1405–1454.

Ciconte, E. (2008). *Storia criminale. La resistibile ascesa della mafia,'ndrangheta e camorra dall'Ottocento ai giorni nostri*. Rubbettino, Soveria Mannelli (CZ).

Commissione Antimafia (2018). Commissione parlamentare di inchiesta sul fenomeno delle mafie e sulle altre associazioni criminali, anche straniere.

Dalla-Chiesa, F., M. Bedetti, F. Cabras, I. Meli, and R. Nicolini (2014). Rapporto trimestrale per la commissione parlamentare di inchiesta sul fenomeno mafioso.

Daniele, G. and G. Dipoppa (2023). Fighting organized crime by targeting their revenue: Screening, mafias, and public funds. *The Journal of Law, Economics, and Organization 39*(3), 722–746.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–2996.

De Simoni, M. (2022). The financial profile of firms infiltrated by organized crime in Italy. *Quaderni dell'antiriciclaggio UIF 17*.

Decarolis, F., R. Fisman, P. Pinotti, and S. Vannutelli (2024). Rules, discretion, and corruption in procurement: Evidence from Italian government contracting. *Journal of Political Economy Microeconomics*.

Dimico, A., A. Isopi, and O. Olsson (2017). Origins of the Sicilian mafia: The market for lemons. *The Journal of Economic History 77*(4), 1083–1115.

DIS (2019). Relazione sulle politica dell'informazione per la sicurezza.

Dixit, A. K. (2004). *Lawlessness and economics: Alternative modes of governance*, Volume 5. Princeton University Press.

DNA (2020). Relazione sulle attività svolte dal procuratore nazionale e dalla direzione nazionale antimafia nel periodo 1° luglio 2018 – 31 dicembre 2019.

Europol (2013). Threat assessment: Italian organized crime.

Fabrizi, M. and A. Parbonetti (2021). The economic consequences of criminal firms.

Fenizia, A. and R. Saggio (2023). Organized crime and economic growth: Evidence from municipalities infiltrated by the mafia. *American Economic Review*.

Fiorentini, G. and S. Peltzman (1997). *The economics of organised crime*. Cambridge University Press.

FT (2024). Organised crime has taken on a different shape in latin america.

Gambetta, D. (1996). *The Sicilian Mafia: The Business of Private Protection*. Harvard University Press.

Gilmore, W. C. (2004). *Dirty money: The evolution of international measures to counter money laundering and the financing of terrorism*, Volume 599. Council of Europe.

Guiso, L., L. Pistaferri, and F. Schivardi (2013). Credit within the firm. *Review of Economic Studies 80*(1), 211–247.

Guiso, L., P. Sapienza, and L. Zingales (2004). The role of social capital in financial development. *American Economic Review 94*(3), 526–556.

Le Moglie, M. and G. Sorrenti (2022). Revealing "mafia inc."? financial crisis, organized crime, and the birth of new enterprises. *Review of Economics and Statistics 104*(1).

Lupo, S. (2009). *History of the Mafia*. Columbia University Press.

Mastrobuoni, G. (2015). The value of connections: Evidence from the Italian-American mafia. *The Economic Journal 125*(586), F256–F288.

Mastrobuoni, G. and E. Patacchini (2012). Organized crime networks: An application of network analysis techniques to the American mafia. *Review of Network Economics 11*(3).

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725.

Melnikov, N., C. Schmidt-Padilla, and M. M. Sviatschi (2020). Gangs, labor mobility and development.

Mirenda, L., S. Mocetti, and L. Rizzica (2022). The economic effects of Mafia: Firm level evidence. *American Economic Review 112*(8), 2748–73.

Parbonetti, A. (2021). La presenza delle mafie nell'economia: profili e modelli operativi. In *La presenza delle mafie nell'economia: profili e modelli operativi: Parbonetti, Antonio*. Padova: Padova University Press.

Pinotti, P. (2015a). The causes and consequences of organised crime: Preliminary evidence across countries. *The Economic Journal 125*(586), F158–F174.

Pinotti, P. (2015b). The economic costs of organised crime: Evidence from Southern Italy. *The Economic Journal 125*(586), F203–F232.

Rexer, J. (2022). The local advantage: Corruption, organized crime, and indigenization in the nigerian oil sector.

Riccardi, M. and P. Reuter (2024). The varieties of money laundering and the determinants of offender choices. *European Journal on Criminal Policy and Research 30*(3).

Schelling, T. C. (1971). What is the business of organized crime? *The American Scholar*, 643–652.

Slutzky, P. and S. Zeume (2024). Organized crime and firms: Evidence from antimafia enforcement actions. *Management Science 70*(10), 6569–6596.

Transcrime (2015). Gli investimenti delle mafie.

Transcrime (2017). Organized crime infiltration of legitimate businesses in europe: A pilot project in five european countries, final report of project ariel.

UIF (2021). *Rapporto Annuale 2021: Unità di Informazione Finanziaria*. Banca d'Italia.

UIF (2022). *Rapporto Annuale 2022: Unità di Informazione Finanziaria*. Banca d'Italia.

UNODC (2011). Estimating illicit financial flows resulting from drug trafficking and other transnational organized crimes.

Varese, F. (2006). How mafias migrate: The case of the 'Ndrangheta in Northern Italy. *Law & Society Review 40*(2), 411–444.

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*.

A Additional Data Sources

We merge the *Mappatura* dataset with several different administrative data, namely the complete register of owners, administrators, and auditors of all firms, firms' operations and balance sheets, credit registry loan-level data, and employment records.

Data on firm composition: owners, directors, and auditors. We observe the identity of owners, directors, and auditors of the universe of firms in Italy using the Infocamere database from the Italian Chamber of Commerce. For each individual, we observe both name and social security identifier, which allow for direct matching with the *Mappatura*. The Infocamere database also provides information on firm location, sector, and years of entry and exit.

Balance sheet and income statement data. We use balance sheet and income statement panel data on the universe of Italian non-financial corporations from Cerved, a standard dataset used in firm-level analyses of Italian firms (Mirenda et al., 2022).[1] The dataset includes operational and financial outcomes such as revenues, payroll, intermediate inputs, assets, liquidity, credit, and debt.

Credit registry. We access loan-level records for all firm-bank credit relationships in Italy through the confidential credit registry database managed by the Bank of Italy.[2]

Social Security aggregates. We use a firm-level panel dataset aggregated from Social Security records (INPS) to study employment counts and average salary at the firm level. Employment and average salary are disaggregated for different worker categories (e.g., managers, white-collar workers, blue-collar workers).

Politicians and lobbying. We obtain data on elected politicians from the Ministry of the Interior. The dataset includes municipal, provincial, and regional-level politicians and national congress members from 1993 to 2023. To merge politicians to owners and administrators, we construct the national identifier of

[1] Firms that are not covered by this data are sole proprietorships or unincorporated partnerships.

[2] We do not observe loans below Euro 75,000 Euro pre-2009 and loans below Euro 30,000 post-2009. See Bofondi et al. (2018) for more details.

the politicians (codice fiscale) based on their demographic characteristics (i.e., full name, age, place of birth). The final dataset includes 575,779 politicians with a national identifier.

For Italian Members of the European Parliament (MEPs), we retrieve data from the European Parliament's Open Data Portal, covering all legislatures since the first election in 1979 until 2019. As with the previous dataset, we compute the national identifier based on demographic information. The final dataset comprises 580 MEPs with corresponding unique identifiers.

We obtain lobbying data by webscraping https://www.lobbyfacts.eu/, a project that, starting in 2012, compiles annual official data on lobbying activities in European institutions, sourced from the EU Transparency Register. For all companies listed in the EU Transparency Register as of May 2024, we retrieve annual lobbying activity data starting from the year they registered. Moreover, for companies that "advance the interests of their clients" (i.e., consultancy firms acting as intermediaries), we extract the yearly list of clients.

To obtain unique firm identifiers, we match the company names involved in lobbying activities with those in the CERVED database using a probabilistic record linkage method provided by De Nederlandsche Bank.[3] Afterward, we retain only matches with a similarity score above 90% and manually review the results, ultimately producing a dataset containing 597 unique identifiers and 227 distinct company names. To overcome the issue of a single company name being matched to multiple identifiers,[4] we then clean the matches by retaining, where possible, only those where the reported headquarters' province, as extracted from the EU Transparency Register, matches the province code contained in the first two digits of the CCIA code, a unique Italian identifier assigned to companies by the Chamber of Commerce.[5] The final dataset comprises yearly observations for 216 unique identifiers, with a total of 818 observations.

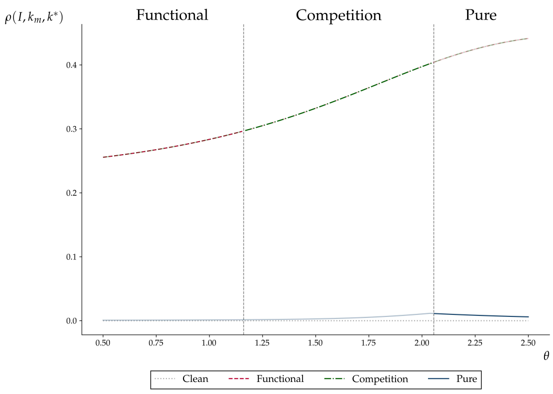
[3] See https://github.com/DeNederlandscheBank/name_matching.

[4] Italian legislation allows two companies to share the same name, provided it does not cause confusion for consumers, meaning they do not compete in the same market.

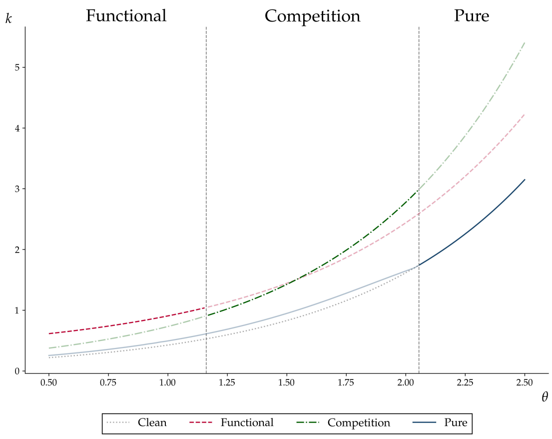
[5] For company names extracted from client lists, we lack information on the headquarters' province. In such cases, we retain only unique matches.

Figure A1: Infiltration Motives & Observable Outcomes (Comparative Statics)

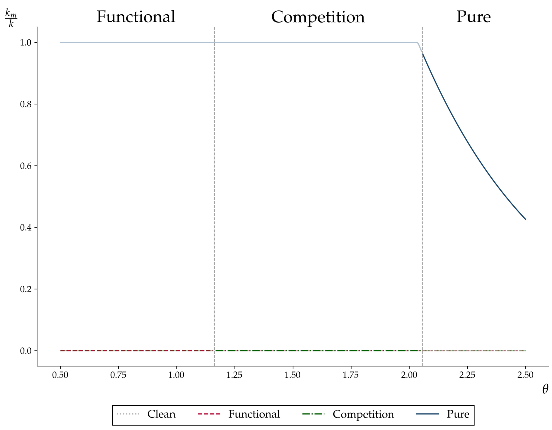
(a) Risk of Confiscation



(b) Risk of Confiscation

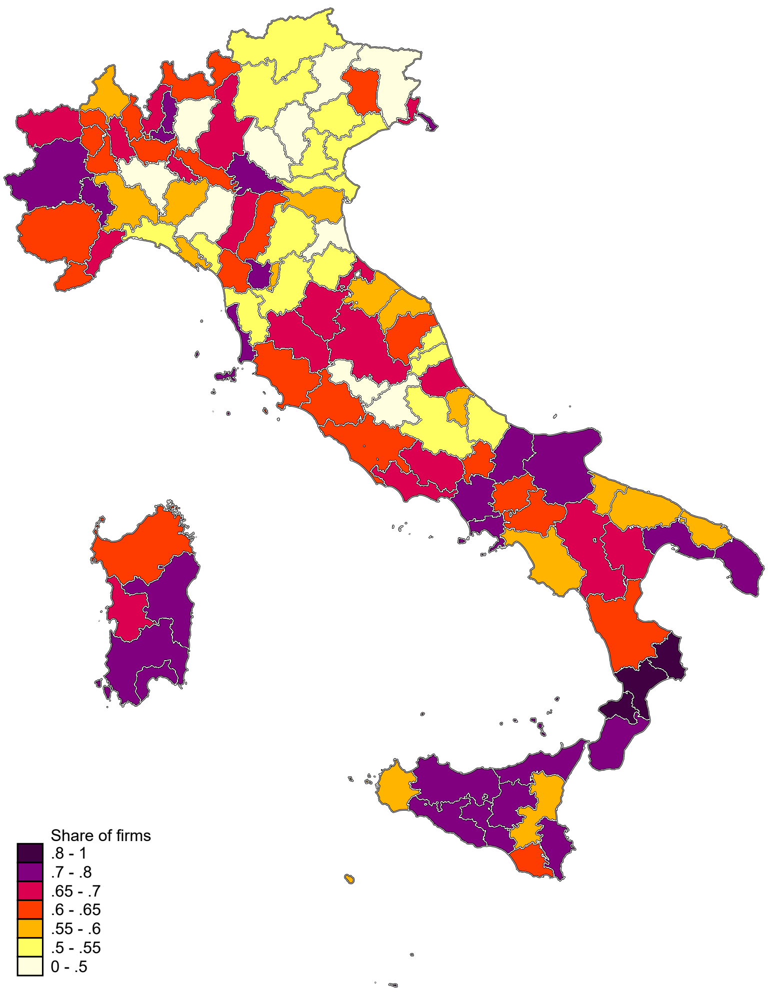


(c) Risk of Confiscation



Notes: The Figure reports observable firms' outcomes – the risk of confiscation $\rho(\cdot)$ (panel (a)), the scale of operation k (panel (b)), and the sources of finance k_m/k (panel (c)) – for different values of θ. Functional forms: $f(k) = \frac{k^\epsilon}{\epsilon}$; $C(k) = \frac{k^\varsigma}{\varsigma}$; $\rho(I, k_m, k^*) = (1 - I)\rho_1(k_m) + I\rho_2(k_m)\rho_2(k^*)$, with ρ_i a logistic function with supremum L_i, growth rate g_i, and midpoint ξ_i. Parameters: $\lambda = \frac{1}{2}$; $r = 0.15$; $i = 0.025$; $\epsilon = \frac{1}{4}$; $\gamma = \frac{1}{3}$; $\varsigma = \frac{1}{3}$ $L_1 = \frac{1}{10}$; $g_1 = 2$; $\xi_1 = e$; $L_2 = \frac{3}{\sqrt{10}}$; $g_2 = \frac{5}{4}$; $\xi_2 = 0$.

Figure A2: Geographic distribution of the share of born-infiltrated firms out of all infiltrated firms

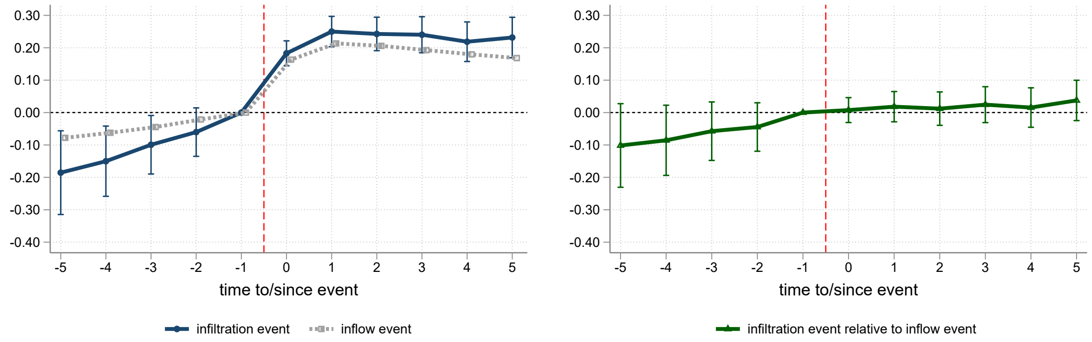


Notes: Geographic distribution of firms potentially connected to organized crime, identified in the *Mappatura*. We present the distribution of the share of infiltrated firms at birth over all infiltrated firms in the province.

Figure A3: Infiltration & Revenues in Mirenda et al. (2022)

(a) Infiltration event vs. inflow event

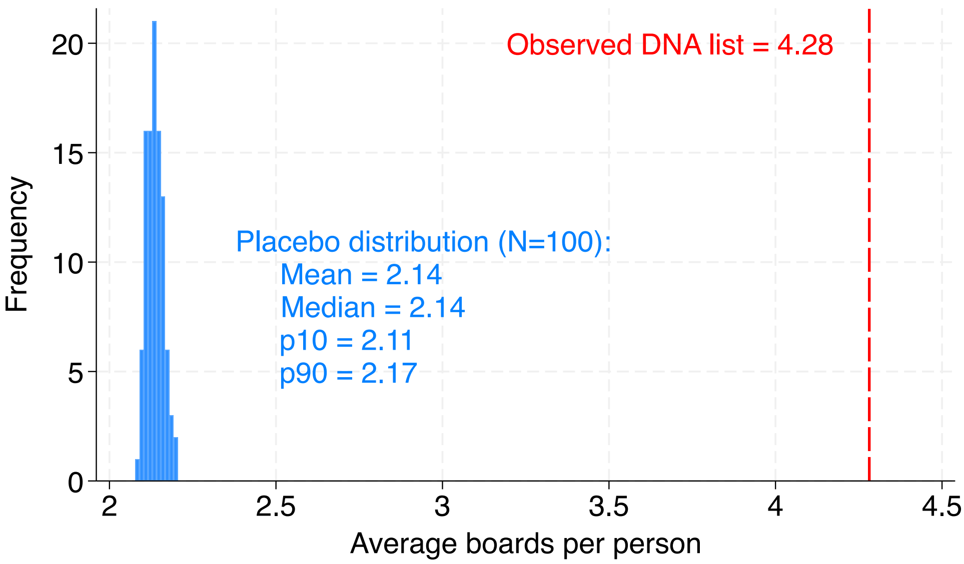
(b) Δ infiltration and inflow



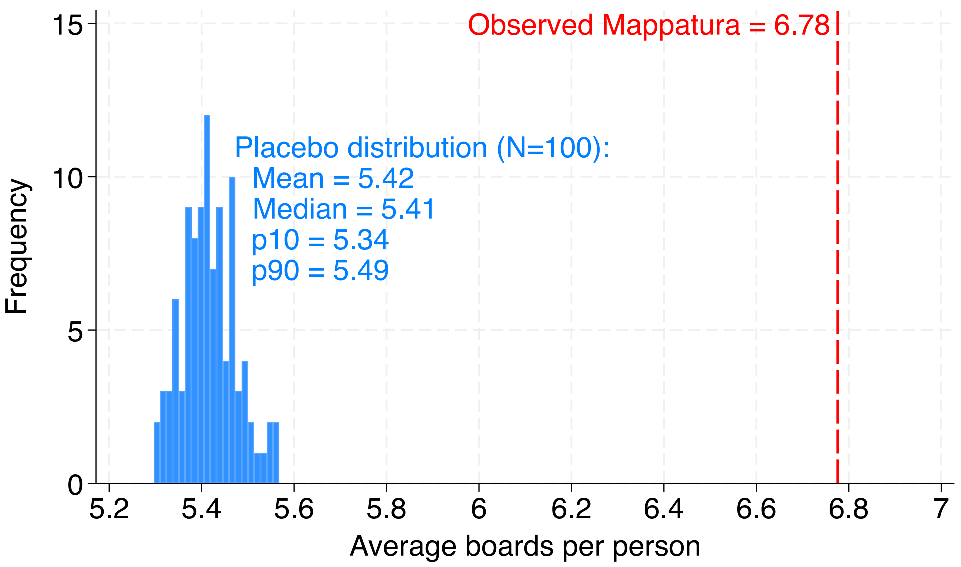
Notes: This Figure replicates our main specification using the data and infiltration definition from Mirenda et al. (2022) and shows that controlling for non-criminal inflows is critical to explain the difference in results. Left panel: Point estimates and 95% confidence intervals for parameters γ_j^{INF} and γ_j^{NC} from equation (6). Right panel: Difference between γ_j^{INF} and γ_j^{NC} estimates. For all firms with an inflow event, we include observations from -5 and +5 years around the event. The specification includes sector-year and province-year fixed effects. Revenues are in the inverse hyperbolic sine form.

Figure A4: Firm links placebo permutation tests

(a) Firm links of individuals of interest to the *DNA*



(b) Firm links of individuals linked to *Mappatura* firms



Notes: We repeat the following exercise 100 times: 1) Draw a random sample of 5,000 owners and 15,000 administrators from the universe of all board members; we treat these 20,000 people as a placebo list of individuals that are of interest to the *DNA* (i.e, Step 1 in the construction of *Mappatura*). 2) Compute the average number of distinct firm links among this placebo *DNA* list. 3) Using these 20,000 individuals, we then create a corresponding placebo list of *Mappatura* firms (i.e, Step 2 in the construction of the true *Mappatura*) and compute the average number of distinct firm links among all individuals who have ties to the placebo *Mappatura*. Panel (a) displays the distribution of the 100 averages of step 2), with the dashed vertical line representing the average among individuals in the true *DNA* list. Panel (b) displays the distribution of the 100 averages of step 3), with the dashed vertical line representing the average among individuals who have ties to the true *Mappatura*.

Table A1: Infiltration and regional characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A	Share of infiltrated firms						
GDP per capita	-0.008***						0.002**
	(0.002)						(0.001)
Financial development		0.002					0.001
		(0.002)					(0.001)
OCGs family names			0.011***				0.016***
			(0.002)				(0.003)
Length court cases				0.010***			0.005**
				(0.003)			(0.002)
Blood donation					-0.006***		0.001
					(0.002)		(0.002)
Trust						-0.008***	0.004
						(0.001)	(0.003)
Mean dep var	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208	0.0208
Panel B	Share of born infiltrated firms						
GDP pc	-0.007*						0.006
	(0.004)						(0.007)
Financial development I		0.006					0.005
		(0.005)					(0.004)
OCGs family names			0.026**				0.014**
			(0.013)				(0.007)
Length court cases				0.015***			0.012*
				(0.005)			(0.006)
Blood donation					-0.005		0.002
					(0.003)		(0.004)
Trust						-0.011**	0.002
						(0.004)	(0.008)
Observations	105	105	86	104	102	105	102
Mean dep var	0.318	0.318	0.318	0.318	0.318	0.318	0.318

Notes: This table presents the correlation between the extent of infiltration and province-level characteristics. In Panel A, the dependent variable is constructed as the number of infiltrated firms alive in each year in a province over the total firms in that province and then we take the average across years. In panel B, the dependent variable is constructed as the number of born infiltrated firms alive in each year in a province over the total number of infiltrated firms in that province and then we take the average across years. *GDP per capita* is the provincial GDP per capita that we average across years. *Financial development* is defined as the variation across firms in the cost at which they can borrow (Guiso et al., 2013). We construct the share of *OCGs family names* by computing the share of people with each last name in mafia home regions (Sicilia, Campania or Calabria), then we keep in each region the top-100 last names. Then, we construct the share of people in non-mafia regions that have any of these last names. The source for the presence of last names is `http://www.gens.info/lib/cog/istruzioni.html`. *Length court cases* is defined as the average length of court cases. *Blood donation* is measured as the incidence of blood donation (Guiso et al., 2004). *Trust* is defined as the average trust on others across different cohorts (Guiso et al., 2004). Robust standard errors are presented in parenthesis. * p<0.1, ** p<0.05, *** p<0.01.

Table A2: Geographic distribution by type of infiltration

	Infiltrated over all firms (1)	Born infiltrated over all firms (2)	Born clean over all firms (3)	Born infiltrated over infiltrated firms (4)
Home region	0.021***	0.004***	0.017***	0.098***
	(0.003)	(0.001)	(0.003)	(0.023)
Observations	107	107	107	107
R-squared	0.579	0.528	0.544	0.145
Mean dep variable	0.0112	0.00369	0.00754	0.626

Notes: This table presents the relationship between infiltration by type and mafia regions. The dependent variable in column 1 (2/3) is the total number of infiltrated (born infiltrated/born clean) firms over the average number of firms in the province, while in column 4, is the share of born infiltrated over all infiltrated firms. *Home region* takes a value one if the firm is located in the provinces of Sicily, Calabria, or Campania. All regressions include year fixed effects. Robust standard errors are presented in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A3: Infiltrated firms and inflow-event firms: Average attributes

	Infiltrated firms		Inflow-event firms		
	Mean	SD	Mean	SD	Diff/SD
	(1)	(2)	(3)	(4)	(5)
Panel A. Geographic location					
North	0.407	0.491	0.494	0.500	-0.176
Center	0.245	0.430	0.257	0.437	-0.028
South	0.348	0.476	0.249	0.432	0.218
Sicily	0.080	0.271	0.051	0.221	0.114
Calabria	0.034	0.180	0.016	0.127	0.110
Campania	0.157	0.364	0.083	0.275	0.231
Panel B. Firm characteristics					
Year of birth	2002.586	12.033	2001.659	12.977	0.074
No. employees	29.130	197.463	10.617	136.479	0.109
(log) Revenues	6.306	2.317	5.708	2.029	0.275
No. managers	2.321	2.331	1.932	1.634	0.193
No. owners	3.351	11.724	2.826	6.364	0.056
Pct. ownership born in Sicily	0.067	0.239	0.048	0.201	0.087
Pct. ownership born in Calabria	0.040	0.183	0.020	0.128	0.123
Pct. ownership born in Campania	0.133	0.323	0.075	0.250	0.200
Panel C. Sectoral composition					
Agriculture, forestry, fishing	0.017	0.129	0.020	0.141	-0.024
Mining, quarrying	0.004	0.059	0.002	0.044	0.031
Manufacturing	0.123	0.328	0.152	0.359	-0.085
Electricity, gas, etc.	0.014	0.119	0.007	0.080	0.077
Water, waste, etc.	0.020	0.139	0.006	0.074	0.126
Construction	0.145	0.353	0.134	0.341	0.032
Wholesale & retail trade	0.188	0.391	0.209	0.407	-0.053
Transportation & storage	0.068	0.252	0.042	0.200	0.117
Accommodation & food services	0.064	0.245	0.067	0.250	-0.011
Information & communication	0.045	0.208	0.054	0.227	-0.040
Finance & insurance	0.009	0.097	0.010	0.098	-0.003
Real estate	0.088	0.284	0.100	0.300	-0.040
Professional business services	0.068	0.251	0.072	0.259	-0.018
Administrative & support	0.075	0.264	0.057	0.232	0.072
Education	0.007	0.082	0.010	0.100	-0.036
Health	0.027	0.162	0.024	0.154	0.016
Arts, entertainment, recreation	0.026	0.159	0.019	0.137	0.046
Others	0.011	0.106	0.015	0.122	-0.033
Number of observations	13,293		697,166		

Notes: Cerved sample, observations with non-zero revenues. Excludes firms born infiltrated. Columns (1) and (2): firm-level means and standard deviations for the year before infiltration. Columns (3) and (4): firm-level means and standard deviations for the year before the inflow event. Column (5): adjusted difference defined as $\frac{\bar{X}_{\text{infiltrated}} - \bar{X}_{\text{inflow}}}{\sqrt{(SD^2_{\text{infiltrated}} + SD^2_{\text{inflow}})/2}}$.

Table A4: Comparison with Mirenda et al. (2022)

	(1)	(2)	(3)	(4)	(5)
	Infiltration: Mirenda et al. (2022)			Infiltration: *Mappatura*	
	Published version	Own construction	Baseline specification	Mirenda et al. (2022) specification	Baseline specification
Post Infiltration	0.237***	0.175***	0.031	0.205***	0.031
	(0.024)	(0.024)	(0.019)	(0.019)	
Post Any Inflow			0.209***		0.206***
			(0.002)		(0.002)
Controls	Yes	Yes	Yes	Yes	Yes
No. observations	6,124,827	9,025,675	7,318,815	8,989,456	7,297,491
No. firms		1154559	1138921	1149005	1133700
No. infiltrated firms		4297	4297	11404	11404
No. inflow firms		-	617104	-	618774

Notes: This table presents the comparison of point estimates between the infiltration definition and research design of Mirenda et al. (2022), and infiltration as defined by *Mappatura* and our research design. In all columns, the estimation sample excludes firms from the South, and include sector-year and province-year fixed effects. The dependent variable is revenue in inverse hyperbolic sine form. In column 1, we report the published estimate from Mirenda et al. (2022). Column 2 shows our results when using their research design and the surnames provided in their replication files. Column 3 estimates equation (5) based on surname-related infiltration. Column 4 uses the *Mappatura* data to identify infiltration, but follows the research design of Mirenda et al. (2022). Column 5 presents our baseline estimates from equation (5). Standard errors clustered at the firm level. * p<0.1, ** p<0.05, *** p<0.01.

Table A5: Born infiltrated and profitability

	(1)	(2)	(3)	(4)
Dep. variable:	Total assets	Exit	Profits >0	Profits/assets
Born infiltrated × Age	0.028***	-0.000	-0.005***	-0.000
	(0.003)	(0.000)	(0.001)	(0.000)
Born infiltrated	0.672***	-0.000**	-0.041***	-0.027***
	(0.016)	(0.000)	(0.003)	(0.001)
Observations	6,126,878	6,126,878	5,600,019	5,591,073
YOB × Province × Industry FE	Yes	Yes	Yes	Yes
Infiltrated firms	22455	22455	20993	20966
Mean dep var	5.875	0.0002	0.697	0.0587

Notes: Born infiltrated is a dummy that takes the value one if the firm was born infiltrated. Age is a continues variable that measure the age of the firm in every year. The sample includes all firms in the CERVED dataset. All regressions include year of birth by province of birth by 2-digit industry fixed effects, and year fixed effects. Standard errors are clustered at the firm level. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Connections of Infiltrat*ed* Firms' Board Members

	Avge. firm connections per person			Avge. infiltrated-firm connections per person		
	(1)	(2)	(3)	(4)	(5)	(6)
=1 if *Mappatura* firm	8.047***	8.055***	7.390***	5.690***	5.691***	5.305***
	(0.144)	(0.144)	(0.156)	(0.074)	(0.074)	(0.075)
Year FE	No	Yes	No	No	Yes	No
Province-sector-year FE	No	No	Yes	No	No	Yes
Mean y, non-*Mappatura*	4.261	4.261	4.215	0.189	0.189	0.181
Mean y, *Mappatura*	12.309	12.309	12.024	5.879	5.879	5.594
No. observations	778,863	778,863	602,672	778,863	778,863	602,672

Notes: Estimates and standard errors of β from the following regression, estimated at the firm level, among the Cerved sample, observations with non-zero revenues, born-clean infiltrated firms and inflow-event firms on the year before infiltration or inflow:

$$y_f = \beta \times 1\{\textit{Mappatura}\text{ firm}\}_f + \psi_{ps,t(f)} + \varepsilon_f,$$

where y_f is either the average firm connections per firm f board member in the year before the event (columns (1)–(3)) or the average infiltrated-firm connections per firm f board member in the year before the event (columns (4)–(6)) and $\psi_{ps,t(f)}$ are province-sector-year fixed effects ($t(f)$ indexes the year before the infiltration/inflow-event of firm f). Robust standard errors clustered at the firm level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Connections of Infiltrat*ing* Individuals

	Firm connections			Connections to infiltrated firms			**1**{Connections to infiltrated firms >0 }		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
=1 OCG-linked person	5.744***	6.395***	1.157***	2.491***	2.346***	0.516***	0.277***	0.279***	0.130***
	(0.696)	(0.603)	(0.042)	(0.481)	(0.325)	(0.018)	(0.006)	(0.005)	(0.002)
Weights	No	No	Yes	No	No	Yes	No	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm $\times$ Year FE	No	No	Yes	No	No	Yes	No	No	Yes
Mean y, non-OCG-linked	7.8	7.8	2.8	0.69	0.69	0.28	0.32	0.32	0.25
No. firms	102,364	92,834	92,834	102,364	92,834	92,834	102,364	92,834	92,834
No. persons	410,351	409,359	409,359	410,351	409,359	409,359	410,351	409,359	409,359
No. observations	4,331,051	4,150,163	4,150,163	4,331,051	4,150,163	4,150,163	4,331,051	4,150,163	4,150,163

Notes: Estimates and standard errors of β from versions of the following regression, estimated among the set of firm-years for which at least one board member (administrator, owner, or auditor) is connected to OCGs:

$$y_{ijt} = \beta \times 1\{\text{OCG-linked}\}_i + \psi_{jt} + X'_{ijt}\gamma + \varepsilon_{ijt},$$

where i indexes people, j indexes firms, and t indexes years. y_{ijt} is either the number of firms that i is connected to in year t (columns (1)–(3)), the number of connections to firms having at least 2 OCG-linked persons that i is connected to in year t (columns (4)–(6)), or a dummy equal to one if the number of connections to firms having at least 2 OCG-linked persons that i is connected to in year t is greater than zero (columns (7)–(9)). In columns (3), (6), and (9), observations are weighted so that each person i receives the same weight. X_{ijt} are controls. Controls in regressions without firm$\times$year FE: dummy for owner, dummy for administrator, firm size (four categories), and year fixed effects. Controls in regressions with firm$\times$year FE: dummies for owner and administrator, interacted with firm size (four categories). Standard errors clustered at the person level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

B Robustness

This Appendix discusses the robustness of our results to three main empirical decisions. First, we discuss the robustness to the empirical design employed in the paper, second, we discuss the robustness to the definition of infiltration, and finally to sample restrictions.

B.1 Empirical design

Our main specification for the "born-clean" analysis relies on a two-way fixed-effects model (TWFE). In the presence of heterogeneous treatment effects, the TWFE model suffers from "bad" comparisons if later treated units are used as a control for early treated units, thus biasing the estimated parameter from the TWFE model. In our context, the number of never-treated units (i.e., firms that did not experience either an infiltration or an inflow event) is substantially larger and this is not a major concern. Following De Chaisemartin and d'Haultfoeuille (2020), the share of the estimates with negative weights (i.e., coming from these "bad" comparisons) are very small (0% for infiltration and 6% for other inflow events). In any case, we perform two robustness exercises for this model.

First, we estimate a stacked-panel regression as in Cengiz et al. (2019). To do this, we create a panel around each cohort that was infiltrated and compare it with the cohort of firms that receive an inflow event. Thus, in this model, we are always comparing infiltrated firms to firms that experienced an inflow event in the same year. In Figure B1 and Table B1, we present the results that are aligned with the main results presented under the TWFE model.

Second, we estimate a static model in the staggered difference-in-difference framework that is robust to the presence of heterogeneous treatment effects. In particular, we estimate the model suggested by Wooldridge (2021). Note that in our context we have two types of firms that are experiencing a staggered change and we are interested in the difference between the two. Therefore, we estimate the coefficient for each group and then test for the difference between the two. Table B2 shows the robustness of our conclusions to implementing this alternative estimation method.

B.2 Measure of infiltration

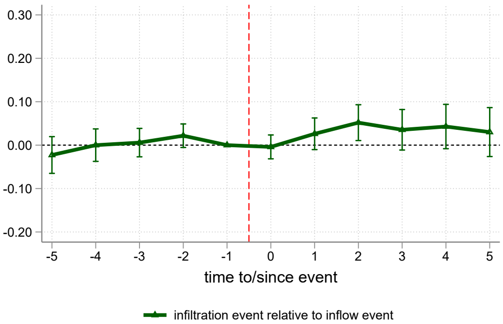
Our analysis relies on the novel and comprehensive data from *Mappatura*. However, there have been other measures used in previous papers to define business-mafia relations. In Tables B3 and B4, we present the robustness of our results to these different measures. In column 1, we present the results for our baseline specification. In column 2, we extend our measure based on *Mappatura*, but we follow a similar strategy as in Mirenda et al. (2022) where we also call infiltrated a firm where the owners of a company faced an infiltration in another firm that they owned. In column 3, we present the measure of infiltration based on surnames by Mirenda et al. (2022). In columns 4 and 5, we present the union of these measures where, in column 4, we use *Mappatura* as in column 1, while, in column 5, we use the extended measure of *Mappatura* as in column 2. Finally, in column 6, to further reduce concerns about false positives, we drop from the sample infiltrated firms that have an risk score equal to 2 (see Section 2.2 for details on the score). Overall, we find our results to be robust to different definitions of infiltration.

B.3 Sample restriction

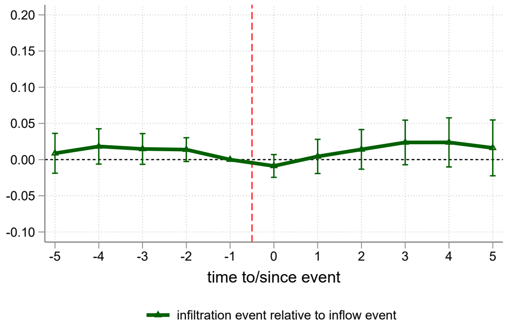
There are two main decisions in terms of sample restrictions that we made in our analysis. The first is that we keep firms with positive revenues for the "born-clean" analysis as in Mirenda et al. (2022). The second is that we estimate the model using the entire country, as opposed to excluding the southern region as in Mirenda et al. (2022). Table B5 presents the robustness to both decisions.

Figure B1: Dynamic specification stacked panel

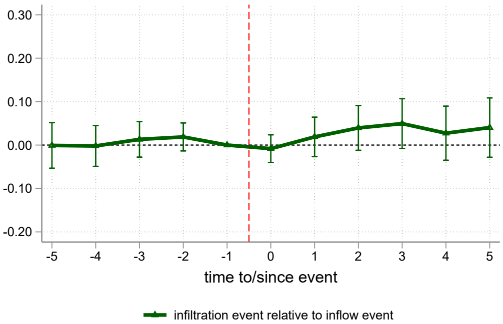
(a) Revenue



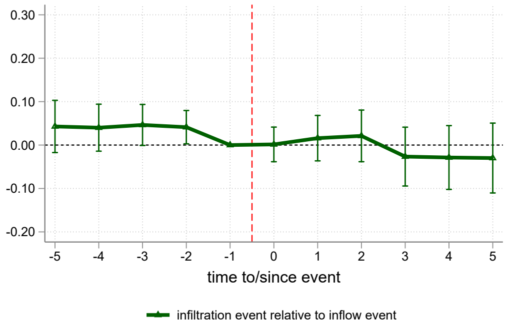
(b) No. Employees



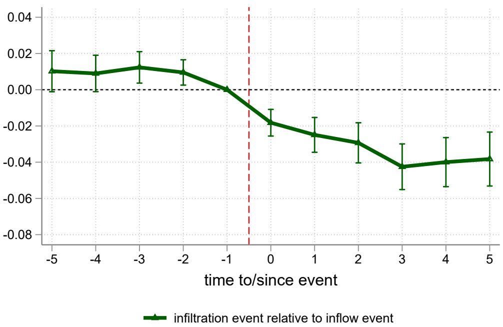
(c) Payroll



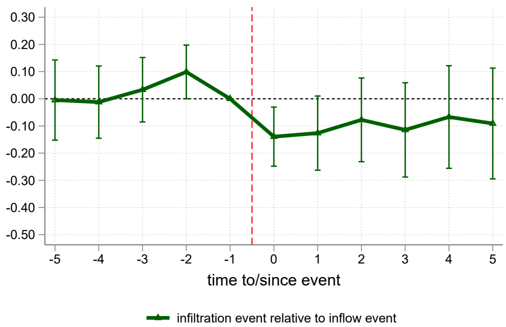
(d) Inputs



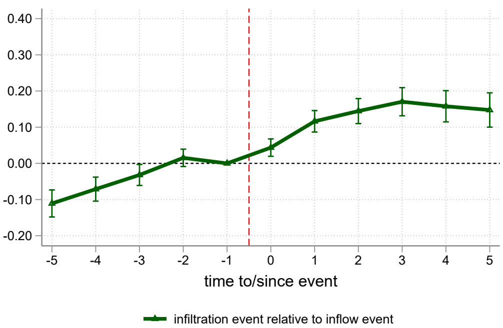
(e) Bank loans, extensive margin



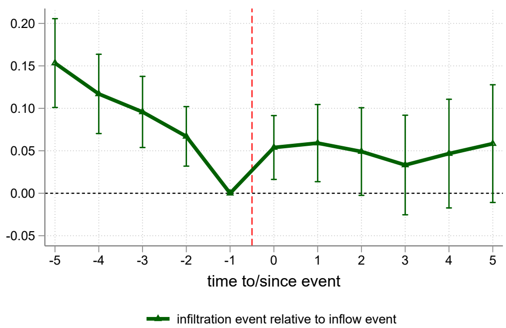
(f) Bank loans, intensive margin



(g) Receivables



(h) Cash holdings



Notes: This figure presents the point estimates and 95% confidence intervals for parameters from a stacked panel specification. The sample includes infiltrated firms and firms that experience an inflow event. The specification includes firm, year-cohort, sector-year-cohort, and province-year-cohort fixed effects. All outcome variables are in the inverse hyperbolic sine, form except for panel (e) which is a dummy.

Table B1: Robustness: Stacked panel regressions

Dep variable:	Revenue	No. Employees	Payroll	Inputs	=1 any bank loans	Bank loans if >0	Receivables	Cash
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Infiltrated × Post	0.020	-0.002	0.014	-0.027	-0.034***	-0.138**	0.132***	-0.009
	(0.015)	(0.010)	(0.019)	(0.022)	(0.004)	(0.059)	(0.013)	(0.018)
Observations	5,382,559	5,382,559	5,382,559	5,382,559	5,382,559	3,100,642	5,224,886	5,382,559
R-squared	0.847	0.904	0.876	0.870	0.732	0.702	0.870	0.693
YOB × Province × Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firms	789795	789795	789795	789795	789795	475124	776959	789795
Infiltrated firms	16788	16788	16788	16788	16788	9831	16533	16788
Mean dep var	6.663	1.622	3.954	4.354	0.524	11.95	5.723	3.724

Notes: This table presents the point estimates from an stacked panel regression. We construct the sample by creating a panel of -5 and +5 years after the treatment event for each cohort of infiltrated firms and firms that experience an inflow event. The sample excludes born-infiltrated firms. $PostInfiltrated$ takes the value one after a firm i is infiltrated, while $Post$ takes the value one after the infiltration or experiece an inflow event. All columns include firm, year-cohort, sector-year-cohort and province-year-cohort fixed effects. All outcome variables are in inverse hyperbolic sine form except for column 5 which is a dummy. Standard errors clustered at the firm level. * p<0.1, ** p<0.05, *** p<0.01.

Table B2: Robustness: Wooldridge (2021)

Dep variable:	Revenue	No. Employees	Payroll	Inputs	=1 any bank loans	Bank loans if >0	Receivables	Cash
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Infiltration								
Post Infiltration	0.272***	0.144***	0.303***	0.227***	0.005	0.120**	0.367***	0.075***
	(0.015)	(0.011)	(0.020)	(0.023)	(0.004)	(0.060)	(0.015)	(0.018)
Panel B: Non-Infiltration inflow								
Post Non-Infiltration inflow	0.213***	0.121***	0.246***	0.205***	0.035***	0.203***	0.180***	0.071***
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.008)	(0.002)	(0.002)
Panel C: Difference								
Difference	0.059	0.023	0.057	0.022	-0.030	-0.083	0.187	0.004
p-value difference	0.000	0.038	0.004	0.342	0.000	0.170	0.000	0.826

Notes: In this table, we present the estimated parameter of interest using the method suggested by Wooldridge (2021) for staggered difference-in-differences. In Panel A, we compare infiltrated to firms that never experience an inflow event, while in Panel B, we compare firms that experience an inflow event to firms that never experience an infiltration. In Panel C, we take the difference of the coefficients and compute the p-value of the difference. * p<0.1, ** p<0.05, *** p<0.01.

Table B3: Robustness: Infiltration definition, operational outcomes

Panel A: Revenues

	(1)	(2)	(3)	(4)	(5)	(6)
Post Infiltration	0.031	-0.038***	0.031	0.028*	-0.031***	0.030
	(0.019)	(0.009)	(0.024)	(0.016)	(0.009)	(0.020)
Post Any Inflow	0.206***	0.204***	0.209***	0.206***	0.204***	0.206***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	7,297,491	7,177,676	7,318,815	7,278,594	7,159,988	7,292,605
No. firms	1,133,700	1,112,625	1,138,921	1,130,820	1,109,966	1,132,801
No. infiltrated firms	11,404	32,179	4,297	15,366	35,482	10,505
No. inflow event firms	618,774	582,656	629,800	613,736	578,439	618,774
Infiltration definition	UIF	Ext. UIF	MMR	UIF ∪ MMR	Ext. UIF ∪ MMR	UIF sidna ≥ 3

Panel B: No. Employees

	(1)	(2)	(3)	(4)	(5)	(6)
Post Infiltration	0.010	-0.008	-0.002	0.005	-0.007	0.007
	(0.013)	(0.006)	(0.016)	(0.011)	(0.006)	(0.014)
Post Any Inflow	0.114***	0.111***	0.116***	0.113***	0.111***	0.114***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	7,297,491	7,177,676	7,318,815	7,278,594	7,159,988	7,292,605
No. firms	1,133,700	1,112,625	1,138,921	1,130,820	1,109,966	1,132,801
No. infiltrated firms	11,404	32,179	4,297	15,366	35,482	10,505
No. inflow event firms	618,774	582,656	629,800	613,736	578,439	618,774
Infiltration definition	UIF	Ext. UIF	MMR	UIF ∪ MMR	Ext. UIF ∪ MMR	UIF sidna ≥ 3

Panel C: Payroll

	(1)	(2)	(3)	(4)	(5)	(6)
Post Infiltration	0.038	-0.031**	0.023	0.031	-0.026**	0.032
	(0.025)	(0.013)	(0.032)	(0.020)	(0.012)	(0.026)
Post Any Inflow	0.236***	0.233***	0.240***	0.236***	0.232***	0.236***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	7,297,491	7,177,676	7,318,815	7,278,594	7,159,988	7,292,605
No. firms	1,133,700	1,112,625	1,138,921	1,130,820	1,109,966	1,132,801
No. infiltrated firms	11,404	32,179	4,297	15,366	35,482	10,505
No. inflow event firms	618,774	582,656	629,800	613,736	578,439	618,774
Infiltration definition	UIF	Ext. UIF	MMR	UIF ∪ MMR	Ext. UIF ∪ MMR	UIF sidna ≥ 3

Panel D: Inputs

	(1)	(2)	(3)	(4)	(5)	(6)
Post Infiltration	-0.027	-0.125***	-0.052	-0.037	-0.118***	-0.032
	(0.028)	(0.014)	(0.035)	(0.023)	(0.013)	(0.030)
Post Any Inflow	0.181***	0.183***	0.184***	0.181***	0.183***	0.181***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	7,297,491	7,177,676	7,318,815	7,278,594	7,159,988	7,292,605
No. firms	1,133,700	1,112,625	1,138,921	1,130,820	1,109,966	1,132,801
No. infiltrated firms	11,404	32,179	4,297	15,366	35,482	10,505
No. inflow event firms	618,774	582,656	629,800	613,736	578,439	618,774
Infiltration definition	UIF	Ext. UIF	MMR	UIF ∪ MMR	Ext. UIF ∪ MMR	UIF sidna ≥ 3

Notes: Point estimates from equation (5) using different measures of infiltration. Column (1), UIF, uses the *Mappatura* definition excluding firms from the South (for comparability with remaining definitions); column (2), Ext. UIF, extends *Mappatura* applying the owners-of-owners procedure (also excluding South); column (3), MMR, uses the infiltration definition of Mirenda et al. (2022); column (4) uses the union of UIF and MMR; column (5) uses the union of Ext. UIF and MMR; column (6) uses *Mappatura* but excludes firms with the lowest risk factor (Sidna=2). The sample excludes born-infiltrated firms. For all firms infiltrated or firms with a clean inflow-event, we include observations from -5 and +5 years after the treatment event. Post Infiltration$_{it}$ takes the value one after a firm i is infiltrated, while Post Any Inflow$_{it}$ takes the value one after firm i is infiltrated or experiences a non-criminal inflow event. All columns include sector-year and province-year fixed effects. All outcome variables are in inverse hyperbolic sine form. Standard errors clustered at the firm level. * p<0.1, ** p<0.05, *** p<0.01.

Table B4: Robustness: Infiltration definition, financial outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: =1 any bank loans			
Post Infiltration	-0.033***	-0.024***	0.001	-0.024***	-0.023***	-0.035***
	(0.005)	(0.003)	(0.007)	(0.004)	(0.002)	(0.005)
Post Any Inflow	0.030***	0.030***	0.030***	0.030***	0.030***	0.030***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	7,297,491	7,177,676	7,318,815	7,278,594	7,159,988	7,292,605
No. firms	1,133,700	1,112,625	1,138,921	1,130,820	1,109,966	1,132,801
No. infiltrated firms	11,404	32,179	4,297	15,366	35,482	10,505
No. inflow event firms	618,774	582,656	629,800	613,736	578,439	618,774
Infiltration definition	UIF	Ext. UIF	MMR	UIF $\cup$ MMR	Ext. UIF $\cup$ MMR	UIF sidna ≥ 3
			Panel B: Receivables			
	(1)	(2)	(3)	(4)	(5)	(6)
Post Infiltration	0.124***	0.024***	0.050**	0.099***	0.027***	0.125***
	(0.018)	(0.009)	(0.024)	(0.015)	(0.009)	(0.019)
Post Any Inflow	0.163***	0.160***	0.166***	0.162***	0.160***	0.163***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	7,297,491	7,177,676	7,318,815	7,278,594	7,159,988	7,292,605
No. firms	1,133,700	1,112,625	1,138,921	1,130,820	1,109,966	1,132,801
No. infiltrated firms	11,404	32,179	4,297	15,366	35,482	10,505
No. inflow event firms	618,774	582,656	629,800	613,736	578,439	618,774
Infiltration definition	UIF	Ext. UIF	MMR	UIF $\cup$ MMR	Ext. UIF $\cup$ MMR	UIF sidna ≥ 3
			Panel C: Cash			
	(1)	(2)	(3)	(4)	(5)	(6)
Post Infiltration	0.006	-0.015	-0.016	-0.001	-0.014	0.008
	(0.023)	(0.012)	(0.030)	(0.019)	(0.012)	(0.024)
Post Any Inflow	0.071***	0.069***	0.071***	0.071***	0.069***	0.071***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	7,297,491	7,177,676	7,318,815	7,278,594	7,159,988	7,292,605
No. firms	1,133,700	1,112,625	1,138,921	1,130,820	1,109,966	1,132,801
No. infiltrated firms	11,404	32,179	4,297	15,366	35,482	10,505
No. inflow event firms	618,774	582,656	629,800	613,736	578,439	618,774
Infiltration definition	UIF	Ext. UIF	MMR	UIF $\cup$ MMR	Ext. UIF $\cup$ MMR	UIF sidna ≥ 3
			Panel D: Bank loans if > 0			
	(1)	(2)	(3)	(4)	(5)	(6)
Post Infiltration	-0.162**	-0.137***	0.055	-0.113*	-0.122***	-0.158**
	(0.072)	(0.033)	(0.091)	(0.058)	(0.031)	(0.074)
Post Any Inflow	0.169***	0.170***	0.170***	0.169***	0.170***	0.169***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. observations	4,201,483	4,129,329	4,211,868	4,192,281	4,120,888	4,198,565
No. firms	654,680	641,556	657,151	653,265	640,290	654,118
No. infiltrated firms	6,950	22,227	2,846	9,582	24,366	6,388
No. inflow event firms	393,124	367,675	399,345	389,975	365,128	393,124
Infiltration definition	UIF	Ext. UIF	MMR	UIF $\cup$ MMR	Ext. UIF $\cup$ MMR	UIF sidna ≥ 3

Notes: Point estimates from equation (5) using different measures of infiltration. Column (1), UIF, uses the *Mappatura* definition excluding firms from the South (for comparability with remaining definitions); column (2), Ext. UIF, extends *Mappatura* applying the owners-of-owners procedure (also excluding South); column (3), MMR, uses the infiltration definition of Mirenda et al. (2022); column (4) uses the union of UIF and MMR; column (5) uses the union of Ext. UIF and MMR; column (6) uses *Mappatura* but excludes firms with the lowest risk factor (Sidna=2). The sample excludes born-infiltrated firms. For all firms either infiltrated or with an inflow-event, we include observations from -5 and +5 years after the treatment event. Post Infiltration$_{it}$ takes the value one after a firm i is infiltrated, while Post Any Inflow$_{it}$ takes the value one after firm i is infiltrated or experiences a non-criminal inflow event. All columns include sector-year and province-year fixed effects. All outcome variables are in inverse hyperbolic sine form except for the dummy variable `=1 any bank loans`. Standard errors clustered at the firm level. * p<0.1, ** p<0.05, *** p<0.01.

Table B5: Robustness: Sample restrictions

Dep variable:	Revenue	No. Employees	Payroll	Inputs	=1 any bank loans	Bank loans if >0	Receivables	Cash
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Exclude 0 revenue restriction								
Post Infiltration	0.026	-0.010	0.003	-0.010	-0.039***	-0.267***	0.205***	0.007
	(0.026)	(0.010)	(0.020)	(0.023)	(0.004)	(0.059)	(0.018)	(0.016)
Post Any Inflow	0.472***	0.148***	0.344***	0.329***	0.041***	0.195***	0.325***	0.141***
	(0.003)	(0.001)	(0.002)	(0.003)	(0.001)	(0.007)	(0.002)	(0.002)
Observations	11,840,215	11,840,215	11,840,215	11,840,215	11,840,215	5,651,835	11,840,215	11,840,215
Mean dep variable	5.284	1.192	3.016	3.537	0.435	11.69	4.787	3.308
Number of infiltrated	20331	20331	20331	20331	20331	10945	20331	20331
Number of inflow firms	925027	925027	925027	925027	925027	516173	925027	925027
Panel B: Exclude firms in the South								
Post Infiltration	0.031	0.010	0.038	-0.027	-0.033***	-0.162**	0.124***	0.006
	(0.019)	(0.013)	(0.025)	(0.028)	(0.005)	(0.072)	(0.018)	(0.023)
Post Any Inflow	0.206***	0.114***	0.236***	0.181***	0.030***	0.169***	0.163***	0.071***
	(0.002)	(0.001)	(0.003)	(0.003)	(0.001)	(0.008)	(0.002)	(0.003)
Observations	7,297,491	7,297,491	7,297,491	7,297,491	7,297,491	4,201,483	7,297,491	7,297,491
Mean dep variable	6.467	1.428	3.583	4.147	0.526	11.80	5.307	3.607
Number of infiltrated	11404	11404	11404	11404	11404	6950	11404	11404
Number of inflow firms	618774	618774	618774	618774	618774	393124	618774	618774

Notes: This table presents the point estimates from equation (5). In panel A, we keep observations with 0 revenue, while in panel B, we exclude firms located in the South. The sample excludes born-infiltrated firms. For all firms either infiltrated or with an inflow-event, we include observations from -5 and +5 years after the treatment event. Post Infiltration$_{it}$ takes the value one after a firm i is infiltrated, while Post Any Inflow$_{it}$ takes the value one after firm i is infiltrated or experiences a non-criminal inflow event. All columns include sector-year and province-year fixed effects. All outcome variables are in inverse hyperbolic sine form. Standard errors clustered at the firm level. * p<0.1, ** p<0.05, *** p<0.01.