Competition and Violence in Illegal Markets: Theory and Evidence from Naples

Davide Zufacchi[*]

05 November 2025

Updated version: click here.

Abstract

This paper studies how market structure–the number and distribution of competitors–shapes the use of violence in illegal drug markets. Using new data on gangs, conflicts, and drug seizures from Naples, I document that fighting reduces the number of active gangs, but it also generates violence, which attracts police attention and disrupts the drug market. I develop and estimate a quantitative model of an oligopolistic market in which consumers choose where to buy drugs, gangs fight to expand market share, and police responds endogenously. The results reveal an inverse U-shaped relationship between market fragmentation and violence, reflecting the combination of a declining probability of conflict with an increasing intensity of each fight. I evaluate three enforcement strategies (kingpin targeting, selective enforcement, and drug legalization), allowing their effects to depend on the initial market structure. This approach, in the spirit of the entry-deterrence literature, reconciles seemingly contradictory evidence on enforcement and violence in illegal markets.

JEL Codes: D74, K14, K42, L13

1. introduction

Illegal drug markets impose substantial costs on society. They absorb large enforcement resources, depress economic activity, and fuel violence (Pinotti, 2015; Monteiro & Rocha, 2017; Alesina et al., 2019;

I am deeply grateful to Salvatore Leotta, Luciano Lombardi, Guido Persico, and Isaia Sales for their collaboration, assistance, and for sharing the data. I am grateful for the suggestions and support of Maurizio Agricola, Alfredo Fabbrocini, Libera D'Angelo, Alessandro Giuliano, Giovanni Leuci. I am indebted to Eliana La Ferrara, Joao Granja, Nicola Persico, Imran Rasul, Tommaso Sonno, and Gabriel Ulyssea for their guidance and support throughout this project. I also thank Christopher Blattman, Javier Boncompte, Marco Castelluccio, Mathieu Couttenier, Magdalena Dominguez, Jeffrey Grogger, Igal Hendel, Jonas Hjort, Gaston Illanes, Rocco Macchiavello, Lars Nesheim, Aureo de Paula, Liyang Sun, and Matteo Vasca for valuable comments and suggestions. I am grateful to seminar participants at the Institute for Fiscal Studies; London School of Economics; NBER Economics of Firearm Markets, Crime, and Gun Violence; Northwestern University; University College London; University of Oxford; the University of Chicago for helpful feedback. Financial support from the Laboratory for Effective Anti-Poverty Policies (LEAP – Bocconi) is gratefully acknowledged. Luca Matarazzo provided excellent research assistance. All errors are mines.

[*]Department of Economics, University College London and IFS, davide.zufacchi.20@ucl.ac.uk

Melnikov et al., 2022; Sviatschi, 2022).[1] Enforcement strategies intended to disrupt these markets have sometimes intensified conflict rather than reduced it (Moeller & Hesse, 2013; Dickenson, 2014; Dell, 2015; Castillo et al., 2020). Such strategies commonly rest on the premise that fragmenting criminal organizations weakens them, thereby reducing violence. Yet this abstracts from the competitive structure of illegal markets, in which gangs interact strategically, often through violence. I propose a new framework to analyze these interactions.

This paper develops an equilibrium model of violent competition in illegal markets to examine how market structure shapes violence and mediates how enforcement strategies affect it.[2] By engaging in conflict, gangs may expand their market share, but they also generate violence that attracts police attention and deters consumers. In equilibrium, the level of violence depends on market structure: on the number of competitors and on the degree of asymmetry among them. An increased number of competitors reduces the marginal market share while expanding the opportunity set of confrontations. Likewise, the degree of asymmetry shifts outcome probabilities but also the marginal cost per fight, due to changes in the violence generated.

To study these trade-offs, I collect new data on gangs, conflicts, and drug seizures in the province of Naples between 2015 and 2022. I begin by documenting novel facts on the use of violence in illegal markets. Building upon these, I develop a new model that incorporates three key elements: (1) a discrete choice demand system in which consumers decide where to purchase drugs, (2) a stochastic fighting game in which gangs compete to expand market share, and (3) endogenous police responses. The unique availability of data on fight outcomes and violence allows identification of key parameters of the fighting technology. I then use the estimated model to quantify how initial market structure shapes violence and to evaluate counterfactual enforcement strategies, allowing their effects to depend on the initial structure, a key novelty of the model.

Four core findings emerge from the analysis. First, the expected number of fights increases with the number of competitors but decreases with the degree of asymmetry between them. This latter result is driven by *diseconomies of scale* in the fighting technology, particularly violence spillovers across locations,

[1] Drug trafficking is linked to nearly half of homicides in the European Union (Eurojust, based on Europol data). The EU's expenditures related to illicit drugs totalled roughly €34 billion in 2013 (EMCDDA, 2013). The direct cost of drug-related homicides alone is €0.5–€5 billion, depending on the statistical value of life (authors' calculations based on EU Policy Department, 2013).

[2] This connects to the classic literature on the industrial organization of crime (Schelling, 1967; Buchanan, 1973; Reuter, 1985; Becker et al., 2006; Levitt, 2017). Relatedly, Castillo et al. (2020) show that violence in Mexico increases with drug scarcity, but only in municipalities where multiple cartels operate.

which disproportionately affect larger gangs. Overall, the relationship between the expected number of homicides and market concentration is inverse U-shaped, reflecting the combination of a *declining probability* of conflict with an *increasing intensity* of each fight. Second, in highly fragmented markets, kingpin strategies (dividing large players) can reduce violence, whereas in concentrated markets such interventions may have the opposite effect. Third, a strong and predictable response to violence (selective enforcement) is generally effective, especially when markets are concentrated. Fourth, the effects of drug legalization on violence depend on the balance between a *negative demand-shock* effect and a *gambling-for-resurrection* effect. Taken together, these results indicate that market structure is central to determining whether a given enforcement strategy mitigates, fails to affect, or unintentionally amplifies violence.

This paper builds on newly assembled data from intelligence, administrative, and survey sources. From intelligence reports, I build matrices of gang presence by location and year, together with detailed records of inter-gang conflicts that identify the parties involved, the year, the contested area, and the outcome. Administrative records on homicides and arrests allow me to measure the consequences of these conflicts. I complement these sources with a new census of approximately 14,000 drug seizures, including information on the city, date, substance type, and quantity seized, as well as a newly collected random sample of 660 chemical analyses used to assess drug purity. Finally, I merge these data with intelligence surveys reporting average retail and wholesale prices, and with additional information on area characteristics.

The data reveal three key features of the fighting technology. First, fights increase gang-related homicides, and this effect intensifies with the scale of the gangs involved: conflicts among small gangs are associated with 0.47 (+337% with respect to peaceful periods) gang homicides on average, whereas those involving at least one large gang are associated with 0.65 homicides (+491%).[3] These fights, in turn, are followed by an increase in arrests. Second, fights reduce the number of active gangs, particularly when large gangs are involved. A fight is associated with a 3% decline in the number of active gangs in the location in the following period, rising to 8% when the fight includes at least one large gang.[4] Third, violence disrupts local drug markets. One additional gang-related homicide is associated with a 12% decrease in

[3]The market is modeled as a set of locations, districts in Naples and cities in the surrounding province, over which gangs compete. A gang's *scale* is defined as the number of locations in which it operates. Throughout the paper, a gang is defined as large if its scale is strictly greater than 3, corresponding to the 90th percentile of the gang scale distribution.

[4]These figures reflect the high share (77%) of fights that end in a draw, meaning that neither of the involved gangs exits the location.

the average quantity of drugs seized.[5] These patterns suggest that gangs face a trade-off when deciding whether to engage in violence: while fighting may force competitors out of locations and expand market share, it also provokes violence which attracts police attention and disrupts market activity.

I build a quantitative model of violent competition and estimate the key parameters governing this trade-off to assess how counterfactual changes in initial market structure affect violence. Consumers choose where to buy drugs by weighing transport costs and local violence within a nested logit framework, following Goldberg (1995). Gangs are heterogeneous in their initial size, which affects both the fighting outcome probabilities and the amount of violence they generate. They play a two-stage game. In the first stage, they decide whether and against whom to fight in a sequential game. In doing so, they balance the gains from increased market share against the losses from reduced demand and heightened police attention, modeled as a fixed cost. This trade-off parallels the literature on predation and entry deterrence, in which firms deter rivals through capacity overinvestment (Spence, 1977; Dixit, 1980), predatory pricing (Milgrom & Roberts, 1982; Benoit, 1984), or other strategic actions (e.g., Schmalensee, 1978; Fudenberg & Tirole, 1986). I embed this logic in static entry models, building on Ciliberto and Tamer (2009) and Aguirregabiria and Ho (2012), to obtain a model of directed competition in which firms may undertake costly actions to expel rivals and capture market share. In the second stage, gangs realize profits given the resulting market structure, demand, and the distribution of police attention.[6]

To simplify the game and reduce the dimensionality of the choice space, I adopt a new dyadic representation. Specifically, the game is viewed as a network of dyads, each corresponding to a pair of gangs and a location, so that the set of dyads spans all possible pairwise combinations of gangs across locations.[7] The model yields a unique prediction for the probability of a fight in each dyad, and I assume that fighting decisions are made independently across them. Formally, each dyad constitutes a separate choice problem, ruling out coordinated strategies that span multiple dyads, although payoffs remain in-

[5] This is the *average*, rather than the total, quantity seized; it is therefore conditional on the number of seizures. Results are unchanged when using the share of total quantity seized and when controlling for the number of seizures to account for changes in enforcement intensity or selection. In the estimation section, I address this endogeneity concern in more detail (please refer to Section 5). Similarly, a fight is associated with a 4.5% decline in the same period and a 13% decline in the following year. Interestingly, I also find an increase in the average purity of drugs during violent periods, consistent with the idea that higher quality is used to attract consumers, followed by a 10% decline after the conflict, in line with weaker competition. The quality results should be interpreted cautiously given the small sample.

[6] Gangs compete only through violence, not on prices or drug purity, to maintain parsimony while preserving the key mechanisms needed to study the strategic use of violence.

[7] A location is a geographical unit over which gangs compete; they are districts in Naples and cities in the surrounding province.

terdependent.[8] This assumption is consistent with limited coordination capacity: gangs may be unable to perfectly synchronize contemporaneous fighting decisions (for example, due to heat-of-the-moment shocks) and instead form expectations about outcomes in other dyads when choosing optimally within each one. Consequently, each dyad can be treated as a player facing a game against nature: given the equilibrium level of violence, the model uniquely determines the probability that fighting is the best response for its members.

I use the newly collected data to estimate the structural model. First, I combine drug seizure records with survey data on prices and administrative data on violence to estimate demand. Two key identification challenges arise: endogeneity from selection into seizures and the endogenous nature of violence. I address the first through two alternative approaches: (i) control for the number of seizures, under a separability assumption, (ii) a classical control-function approach (Heckman, 1979) using the share of arrests by the local police as an excluded instrument.[9] The second is addressed using a new instrumental variable for violence: the release of gang members from prison. Releases shift the propensity of gangs to engage in conflict without directly affecting demand, effectively serving as a (perceived) cost shifter. Second, I infer the parameters governing the fighting technology from observed fight results, a unique feature of the data. Variation in arrests following homicides, and in homicides following fights, identifies parameters related to police reactions and violence determination. Finally, I recover the fixed costs of police enforcement faced by gangs, as well as the parameters governing unobserved fighting preferences, using a nested fixed-point algorithm. The estimated model closely matches empirical correlations in the data.

Leveraging the estimated model, I study how counterfactual changes in initial market structure affect violence. I vary the number and distribution of gangs by randomly merging or splitting existing gangs and adding or removing gangs, changing the initial ownership matrix. For each structure, I solve a fixed-point algorithm to compute the equilibrium of the fighting game. Simulations show that the expected number of fights increases with the number of competitors, holding their size distribution constant, and decreases with size asymmetry among them, holding their number constant. This result reflects diseconomies of

[8]Importantly, this assumption does not exclude economies of scale in the fighting technology, a key mechanism of the model, since the probabilities of fighting outcomes are heterogeneous across gang sizes. It does, however, restrict cross-location strategies, such as retaliation or deterrence through the threat of fighting in other territories. This limitation is unlikely to be significant in markets such as Naples, where the probability that a pair of gangs operates in only one location is high (74%), but it may be more relevant in contexts with wider spatial overlap among competitors.

[9]Unlike arrests for drug-related crimes, which may directly affect the drug market, local police arrests are relevant for seizures but unlikely to influence drug demand. As discussed in Section 2, this force focuses primarily on traffic and urban enforcement, and drug use is not classified as a major crime in Italy.

scale in the fighting technology outweighing the higher winning probabilities. Violence spillovers play a key role in generating these diseconomies. Summarizing both dimensions in an index of market fragmentation, I find a negative relationship between market concentration and the expected number of fights, and an inverse-U-shaped relationship between market concentration expected gang homicides. This arises from the combination of a *declining probability* of fights and an *increasing intensity* of each fight. Consequently, although the overall level of violence is similar at the two extremes of fragmentation, its nature differs: fragmented markets exhibit frequent, low-intensity fights, whereas concentrated markets feature infrequent, high-intensity conflicts.

These findings shed light on why kingpin strategies (fragmenting larger players) can unintentionally increase violence. In fragmented markets, further fragmentation tends to reduce violence. However, in more concentrated, oligopolistic markets, dismantling large criminal organizations can shift the system from a relatively peaceful equilibrium to a more violent one. In my simulations, a 1% decrease in concentration around the middle of the distribution increases expected gang homicides by about 5.5%.

To conclude, I evaluate two additional enforcement strategies: (i) selective enforcement, modeled as a stronger police response to violence, and (ii) drug legalization. The simulations show that a tougher police response lowers the probability of fights, especially in moderately concentrated markets. Thus, in these markets, selective enforcement can be more effective than fragmenting gangs.[10] For the second strategy, I model legalization as setting the illicit price to zero for some drugs, thereby collapsing the illegal market. Without fixed costs, this acts as a *negative demand shock*, reducing the returns to additional market share, and unambiguously lowers violence. With fixed costs, a *gambling-for-resurrection* effect emerges: gangs may engage in more aggressive competition to survive. At low concentration, the profitability effect dominates and violence falls; at higher concentration, the gambling effect can prevail, implying that legalization may unintentionally increase violence.

Literature and contribution. This paper contributes to the literature on the industrial organization of crime and enforcement. A first strand studies how criminal groups organize, compete, and sustain cooperation in the absence of legal institutions, focusing on internal hierarchies, reputational mechanisms, and governance structures (Polo, 1995; Levitt & Venkatesh, 2000; Mastrobuoni & Patac-

[10] This finding is consistent with Papachristos and Kirk (2015), who report a 23% reduction in shootings following the implementation of Chicago's VRS program.

chini, 2012; Campaniello et al., 2016). Other work examines the determinants of market structure in illicit sectors, emphasizing transaction costs (Cook et al., 2007), the absence of regulatory institutions (Schelling, 1967; Donohue & Levitt, 1998), and moral hazard and search frictions (Galenianos et al., 2012; Galenianos & Gavazza, 2017). Relatedly, Gambetta & Reuter, 2017, Le Moglie & Sorrenti, 2022, and Mirenda et al., 2022 study how criminal organizations influence legal economic activity. This paper contributes to this literature by combining new data with a structural framework to study violent competition in illegal markets. The model captures the strategic interactions among gangs, allowing the analysis of how violence both shapes and responds to market structure and enforcement activities. Moreover, the same framework can be applied to study other forms of targeted negative actions beyond violence, such as negative advertising (Grossman & Shapiro, 1984; Bass et al., 2005; Bostanci et al., 2023) or political confrontation (Coate, 2004; Schultz, 2007; Barton et al., 2016).

A second strand examines the causes and consequences of violent competition in illegal markets. Beginning with Levitt and Venkatesh (2000), this literature highlights that violence in illegal markets responds to economic incentives. Angrist and Kugler (2008) show that the expansion of the cocaine trade increased violence in Colombia, whereas Lind et al. (2014) find the reverse causal link for opium cultivation in Afghanistan. Brown et al. (2025) show that a truce among gangs in El Salvador reduced violence but increased extortion, while Blattman et al. (2024) document that Colombian gangs restrain external violence to protect drug rents from police activity. In Mexico, Dell (2015) find that crackdowns on incumbent cartels increased violence, and Castillo et al. (2020) show that violence rises with drug scarcity, particularly where multiple cartels compete. This paper builds on this literature by exploring the mechanisms underlying these reduced-form results and by providing new evidence on how violence depends on the number and distribution of gangs in the market.

A related body of work evaluates enforcement strategies. Selective enforcement, involving sustained interventions in specific areas, is linked to reduced violence in some settings (Corsaro et al., 2012; Fox et al., 2015) but little to none in others (Klement & Blokland, 2023; Bhatt et al., 2024). Kingpin strategies, which target high-ranking members of criminal organizations, often generate unintended consequences, with many studies finding increased violence (Moeller & Hesse, 2013; Dickenson, 2014; Dell, 2015; Lindo & Padilla-Romo, 2018; Cruz & Torrens, 2023), while others identify cases where they reduce it (Vargas, 2014; Phillips, 2015; Burke, 2023). This paper builds a structural model in which counterfactual enforcement interventions endogenously depend on the initial market structure, reconciling seemingly contradictory evidence on the effects of different enforcement strategies on violence.

The paper is structured as follows. Section 2 introduces the setting of the study and the data. In section 3, we present the facts on the market and present the trade-off between market share and profit margins. Section 4 builds the stuctural model. Section 5 discusses identification and estimation of model parameters. These are used to construct counterfactual scenarios in Section 6. Section 7 concludes.

2. SETTING AND DATA

2.1. *Organized crime in the province of Naples*

Organized crime has a long history in the province of Naples, with some evidence of gang behavior going back to the 17th century, and the first written proof of its existence dating back to the 1800s (Barbagallo, 2014; Sales, 2022). The historian Marc Monnier describes these gangs as "a kind of schismatic police" that controlled Naples' prisons, markets, and illegal gambling houses (Monnier, 1862). These organizations experienced a significant evolution in the post-World War II period, transforming from a local phenomenon into an international criminal network involved in cigarette smuggling, drug trafficking, and other illicit activities (Sales, 1994, 2022). These new opportunities, along with business from the post-earthquake reconstruction following the 1980 disaster, provided the gangs with immense financial resources. At the same time, they created incentives for heightened competition for dominance among rival groups, leading to the first large gang fight in 1980 (Marrazzo, 2005), which resulted in approximately 1,000 homicides in six years.

From the 1990s to today, organized crime in Naples has maintained its horizontal structure with multiple independent gangs (Brancaccio & Martone, 2019). These clans are usually smaller than other Italian mafia groups and are characterized by lower internal organization and more intense conflicts (Catino, 2014). These characteristics make most of these gangs similar to some in North and South America (Dugato et al., 2020). This criminal ecosystem is structured around two dominant cartels. While these two cartels try to maintain control over distinct macro-areas of metropolitan Naples, they are mainly composed of a complex network of small satellite groups that exercise tight territorial control over specific neighborhoods and districts. Therefore, this dichotomous structure is marked by constant instability and fluid dynamics. Even within areas of exclusive cartel control, precarious equilibria persist due to frictions between gangs–both across and within cartels–and the emergence of new criminal groups (for examples of these frictions, please refer to reports from DIA, 2015 to DIA, 2022). As a result, this ecosystem is extremely violent, with mafia-related homicides accounting for 60-70% of all intentional homicides in its territories, significantly exceeding the national average of 10-15% (Catino, 2014). The scale of this "system" is massive, with activities in numerous countries (DIA, 2020), and thousands of affiliates divided into more than 200 gangs.

Today, these gangs engage in a wide range of criminal enterprises, with their core activities including drug trafficking, extortion, infiltration of public contracts, arms trafficking, counterfeiting, usury, protection rackets, prostitution, illegal waste disposal, and money laundering (Barbagallo, 2014). Estimated revenues from all these activities amount to approximately 9 billion euros annually (Transcrime, 2013). Despite this wide diversification, drug trafficking remains their main area of activity, accounting for more than 50% of estimated revenues (Eurispes, 2005). This is consistent with other criminal groups worldwide, as the drug business is often recognized as the most lucrative market for criminal organizations (Reuter, 2014).

Neapolitan gangs involved in the drug market tend to control the retail market by enlisting low-level criminal labor. Direct management of drug dealers creates an efficient distribution system and generally adheres to precise territorial boundaries with other gangs (Becucci, 2004). The exact production and distribution process of drugs varies dramatically among different gangs. However, it develops around some common grounds. International brokers, allied with the previously mentioned cartels, or operating independently, usually import raw drugs from other continents (mostly from Africa and South America - DCSA, 2024). When the broker arrives with the delivery, each gang takes its share, then cuts, packages, and prepares the drug for the retail market (Gribaudi et al., 2009). Local drug dealers then sell the drugs on the street, or, more recently, online. There are defined and organized areas dedicated to the sales of illegal drugs, and their control has often fueled fights between gangs, resulting in numerous casualties (Catino, 2019). The management of dealers is very heterogeneous: some receive a fixed salary or a daily payment based on sales, while others operate as independent entrepreneurs. As well as paying salaries, a gang's common fund is often used to pay for assistance and legal advice, the "tools of the trade" (weapons, cars, phones, etc.), and health and housing emergencies that members may face (Catino, 2019).

As an example of how well organized retail drug distribution is, here we cite one of these dealers (Ferrillo, M., *Spacciatore o spacciato?* Torino: Gruppo Abele (unpublished 2000, p. 39), from Becucci, 2004):

[...] street dealers are called to meet in a large apartment where drugs are distributed along with instructions.

Everybody is assigned a precise spot to stand and somebody warns the dealer right away in case of danger when the police are on the way.

As a result, we should think about the retail drug market in the province of Naples as *violent*, *well-organized*, and *territorial*.

2.2. Italian enforcement system

Before turning to the actual data, it is worth discussing in a few words the organization of the Italian law enforcement system. This system operates through four main police forces, each with distinct responsibilities and jurisdictions. The *Polizia di Stato* (State Police), under the Ministry of Interior, primarily handles urban policing, immigration, highway patrol, and railway security, operating mainly in cities and organized in provincial departments. Therefore, the whole Neapolitan province will have only one State Police division. The *Arma dei Carabinieri*, a military force with police duties under both the Ministry of Defense and Interior, serves as a national gendarmerie force, particularly active in rural areas and small towns. The *Guardia di Finanza*, under the Ministry of Economy and Finance, specializes in financial crime, tax evasion, smuggling, and international drug trafficking. Finally, the *Polizia Locale* (Municipal Police), managed by individual municipalities, focuses on local law enforcement, traffic control, and urban regulations enforcement within specific city boundaries. These forces collaborate through multiple inter-agency organizations. For example, the Direzione Investigativa Antimafia (DIA) coordinates anti-mafia operations across all police forces, while the Direzione Centrale per i Servizi Antidroga (DCSA) specializes in coordinating drug-related investigations and operations among the various police forces, facilitating information sharing and joint operations against drug trafficking. We use data coming from all these agencies; often together, sometimes separately.

2.3. Data and definitions

This paper draws on a new data set on the retail drug market, constructed by combining intelligence, administrative, and survey sources.

A key consideration is the data set's geographical granularity. Administrative data are available only at the city–year level. For instance, we observe the annual number of arrests for each city in the province, which implies a single value for the entire city of Naples–a limitation inherent to the nature of the data. By contrast, intelligence data are collected at a finer scale: city–year for locations outside Naples and district–year within Naples city, which consists of 30 districts. This greater granularity enables more detailed geographical analysis, allowing us to track outcomes such as homicides and gang presence at the district level within Naples.

The sample covers all districts and cities in the province of Naples from 2015 through 2022.

Definition 1. *A **location** is defined as either a district within the city of Naples or a city in the province. This is the geographical level at which fights are determined.*

Definition 2. *A **market** is defined as the collection of all locations in the province of Naples for a given drug in a given year.*[11]

Gangs presence Matrices that record gang presence across locations and years are constructed from intelligence sources in two steps. First, DIA reports (e.g., DIA, 2020) are used to build a matrix indicating whether a given gang is mentioned in a specific area during the year under study. Mentions may appear either in maps (e.g., Figure A1) or in the report text (e.g., Figure A2 illustrates the entry of a gang into a new district).[12][13] Then, these matrices are submitted to the Neapolitan police, which confirmed their validity. This process produces eight matrices, one for each year in the sample, reflecting the presence of gangs across locations.

Definition 3. ***Gang scale*** *is defined as the number of locations in which a gang is present during a given year.*

Definition 4. *A gang is defined as **large** if its scale in the period is higher than 3, the 90th percentile of the gang scale distribution.*

Fight matrices The construction of the fighting matrices follows the same procedure as that used for the gang–presence matrices. A list of fights is compiled from DIA reports, recording the parties involved, the year, the contested locations, and the winner.

Definition 5. *Two gangs are defined as having a **fight** for a given location in a given year if DIA reports indicate that they were in conflict over that territory and/or that frictions between them were registered for activities in that territory.*

[11] We use this definition of market because this is the level of granularity at which we observe prices.

[12] The maps are only indicative of gang presence. For this reason, we turn to the text. Analyzing it, we reconstruct the matrix of presence based on both sources.

[13] Throughout data construction, a gang is defined as group of individuals mentioned in the reports under a single name, which is clearly not a subgroup of another gang. While this definition has limitations, the absence of a comprehensive gang list makes it, to our knowledge, the only feasible approach. Accordingly, the gang name reported by the DIA is used as the identifier.

Definition 6. *A gang is defined as* ***winning*** *a fight for a given location if, in the subsequent period, the opposing gang exits the contested location. If no gang exits, the fight is recorded as a draw for that period.*[14]
Then, the lists are submitted to the Neapolitan police, which confirmed its validity.

Drug seizures I use administrative data on drug seizures in the province of Naples, collected by the Italian antidrug authority (Direzione Centrale Servizi Antidroga). Enforcement agencies are required to report each seizure characteristics, including city, date, enforcement authority, type (territorial or maritime; only territorial seizures are retained), substance type, and quantity seized in kilograms. The data set covers approximately 14,000 seizures during the sample period. The analysis focuses on three primary substances: marijuana, heroin, and cocaine. Seizures of these drugs are included.[15] Only seizures with non-missing and positive quantities are used, as quantities are sometimes recorded as zero when negligible amounts are confiscated. For each city, year, and drug, total quantity seized, average seizure size, and number of seizures are computed.

Other data The data set is complemented with additional data from different sources, including intelligence data on the chemical composition of seized drugs, survey data on drug consumption and prices, survey data on city characteristics (including the activity of municipal police), and data on distances between geographical units. A detailed description of these sources is provided in Appendix D.

Data quality The paper combines administrative and intelligence data. Assuming completeness and accuracy of the administrative data, the quality of the constructed intelligence data is assessed in Appendix E by examining how variation in administrative data corresponds to changes in the intelligence outcomes.

3. DESCRIPTIVE EVIDENCE

This paper studies violence as a strategic action of firms operating in illegal markets and examines its relationship with market structure. This section presents preliminary descriptive evidence. First, I provide

[14]A potential inconsistency arises when DIA reports state that a gang (or group of gangs) has won a fight and explicitly mention that the opposing gang left the area–or similar–but the same opposing gang is still recorded as present in the following period. These cases are treated as follows: (1) the (group of) gang is recorded as having won the fight, (2) the opposing gang is assumed to have exited the area, and (3) the subsequent presence of the opposing gang is interpreted as re-entry.

[15]Specifically, seizures classified as "cocaine", "crack cocaine", and "liquid cocaine" for cocaine; "heroin" for heroin; and "hashish", "marijuana", and "cannabis plants" for marijuana are retained.

summary statistics on gangs and their presence across locations over time. Second, I document correlations between the probability of fights and both the number and characteristics of locations. Third, I present new facts on the fighting technology and, building on these, discuss the trade-off gangs face when deciding whether to engage in violence. These patterns motivate the structural model developed in Section 4.

3.1. Summary statistics

Table 1, Panel A, reports summary statistics for city–year averages of key market characteristics. There are several figures worth pointing out. First, gangs are widespread, with an average presence of 72%. Second, fights between gangs are relatively frequent, with an average probability of 7%. Relatedly, the share of gang homicides is high, accounting to approximately 60% of all homicides. Third, the retail drug market is large: more than 200 individuals are arrested on average each year for drug-related crimes, and seizures can reach up to 4,000 kilograms of cocaine in a single year. Fourth, retail prices are consistently above wholesale prices for all drugs and periods. Finally, average drug quality–measured as the percentage of pure substance–varies substantially across city–years, with even greater variation across seizures, as documented in Appendix E.

Table 1, Panel B, reports summary statistics on fights between gangs. In total, 128 fights are observed during the sample period, with the contested location identified in 122 cases. A winner is known in 43 of these. Fights typically involve two *groups* of gangs, with an average of 2.87 gangs participating; the minimum is 2, while the maximum is 14. Only a small fraction of fights (1.6%) involve gangs that were not present in the contested location at the time of the conflict. By contrast, in most cases (82%), both gangs were active in the location they fought over. Finally, the majority of fights (76%) last only one year, with an average duration of 1.319 years. In what follows, we restrict attention to fights in which the contested location is known, accounting for 95% of the sample.

Figure A3 presents the distribution of the number of active gangs across locations. About 20% of location–years record no active gang, implying that nearly 80% have at least one.[16] Locations with only one active gang are relatively rare, accounting for 25% of observations. In the remaining 55%, multiple gangs coexist, with as many as ten active in the same location during a given year. Figure A4 shows the distribution of gang scale, defined as the number of locations in which a gang operates at a given time.

[16] This figure differs from that in Table 1, where the unit of observation is the city–year rather than the location–year.

Table 1: Summary statistics

Panel A: summary statistics sample	Mean	SD	Min	Max	N
Gangs					
Presence	0.72	0.45	0	1	736
Number gangs	3	12.81	0	134	736
Max scale	4.38	2.23	1	9	528
Mean scale	3.29	1.64	1	7	528
Violence					
Fight	0.07	0.25	0	1	736
Number attempted homicides	0.82	7.94	0	156	736
Number homicides	0.41	2.26	0	36	736
Number gang homicides	0.24	1.57	0	28	736
Enforcement activity					
Arrests for drugs	206.90	1,918.97	0	21,180	736
Arrests local police	15.13	217.46	0	5,340	736
Number cocaine seizures	5.04	23.01	0	239	736
Number heroin seizures	1.03	7.81	0	105	736
Number marijuana seizures	10.45	50.09	0	546	736
Drugs					
Quantity seized cocaine (kg)	2.44	20.72	0	490.06	736
Quantity seized marijuana (kg)	27.56	324.42	0	8,435.65	736
Quantity seized heroin (kg)	1.14	17.83	0	466.86	736
Retail price cocaine (euro)	77.81	10.50	67.50	95	736
Retail price heroin (euro)	63.75	9.93	45	75	736
Retail price marijuana (euro)	11.19	2.09	9	15	736
Wholesale price cocaine (euro)	38.25	2.44	35	42	736
Wholesale price heroin (euro)	25.19	5.30	15	32	736
Wholesale price marijuana (euro)	2.59	0.54	2	3.60	736
Average quality cocaine (% pure substance)	63.25	13.98	25.06	84.50	40
Average quality heroin (% pure substance)	19.60	10.28	0.63	38	19
Average quality marijuana (% pure substance)	12.16	5.67	0.20	29.21	57

Panel B: summary statistics fights	
# fights	128
# fights with known contended location	122
# fights with (known) winner	43
Mean number of gangs involved	2.87
Min number of gangs involved	2
Max number of gangs involved	14
# fights in areas with no gang present	1.6%
# fights in areas with one gang present	16.4%
# fights in areas with both gang present	82%
Average number of years in fight	1.319
Percentage fights lasting one year	76%

Notes panel A: This panel displays summary statistics for selected variables in the data. Each observation corresponds to a city year. The number of observations in max scale and mean scale is lower because these variables are missing in cities/years with no gangs ($736 \times 0.72 \approx 528$).
Notes panel B: This panel displays summary statistics on fights. Fighting matrices (including information on the actors, contended area, year, and winner) are constructed using intelligence data from the DIA reports. More information in section 2.3.

Roughly 60% of gangs are active in a single location. The distribution, however, has a long right tail: around 20% of gangs operate in two locations, while the remaining 20% are active in three or more. [17]

Figure A6 presents the cross-sectional variation in the number and average scale of gangs. A key insight is the absence of systematic sorting, indicating substantial cross-sectional variation: some locations have a single small gang, others a single large gang, and others a mix of gangs of different scales. Figure A7 presents the transition matrix for the number of active gangs in a given location. Transition probabilities are strictly positive outside the main diagonal, with the largest values concentrated in locations with a moderate number of gangs rather than in those with none or many. This indicates significant temporal variation in the number of gangs. Figure A8 illustrates variation in average gang scale over time. In the absence of such variation, all observations would lie on the 45-degree line, which is clearly not the case. In fact, average scale rises in some locations and falls in others, generating a dispersed distribution.

3.2. Fights and location characteristics

Table A1 presents the correlations between the probability of observing a fight and several characteristics of the location. The probability of conflict is positively correlated with the number of gangs, with an additional gang being associated with a threefold increase in the probability of fights relative to locations with only one gang. Conversely, the average gang size is negatively associated with the probability of conflict, with a one-unit increase in average size corresponding to a 25% decrease in the probability of fights. A similar negative correlation is observed between the probability of fights and both the standard deviation of gang sizes within a location (-22%) and the maximum size observed (-8%). Consistent with basic economic reasoning, a higher share of consumers is positively associated with the probability of fights, with a 1% increase in the share corresponding to a 0.5% increase in the probability of conflict. Finally, the more isolated a location is–hence, the greater its average distance from other locations–the lower the probability of conflict observed (a 1% increase in distance is associated with a 0.3% reduction in the probability of conflict).

To examine more closely the correlations with average gang scale, Table A2, Panel A, reports how these correlations vary across different outcome variables, control specifications, and fixed effects. As before, the probability of conflict is positively correlated with the number of gangs and negatively correlated with

[17] Although some gangs operate in multiple locations, their spatial presence does not display spatial autocorrelation. Figure A5 reports the distribution of Moran's I index for each gang. For single–location gangs, the index is trivially zero. For multi–location gangs, the distribution is centered near zero with a slight positive tail, but overall remains far from the 0.5 threshold typically associated with spatial autocorrelation.

gang size, even when both variables are included jointly (column 3). Columns 4–6 add location and year fixed effects, and the conclusions remain unchanged. Panel B replicates the analysis using the maximum rather than the average gang size. Panels C–E extend the specification to homicide outcomes: the probability of any homicide (Panel C), gang-related homicide (Panel D), and the number of gang homicides per 1,000 inhabitants (Panel E). Panels F and G replicate the analysis for for probability and number of non-gang homicides. Interestingly, no similar trends emerge for the number of gang homicides; I revisit this insight in the counterfactual analysis.[18]

Appendix E further examines these correlations, studying gang entry with event studies and an IV strategy.

These results support the idea of violence as a strategic action in illegal markets. The correlations presented are later used to test the model.[19]

3.3. Facts on the fighting technology

Fact 1: Fights increase violence, especially when large gangs are involved

Table A4 report results from regressions of fight indicators on (i) the number of gang homicides (Panel A) and (ii) arrests for both drug related crimes and conspiracy (Panel B).[20] Columns 1-3 and 11-13 include location fixed effects, while columns 4-10 and 14-20 add year fixed effects. Columns 7, 9, 17, and 19 drop observations with fights involving large gangs, and columns 8, 10, 18, and 20 drop those involving only small gangs. In columns 11–20, the fight indicator equals one both during and after an outbreak.

As expected–and serving as a data-quality check (as in Appendix E)–the average number of gang homicides is higher with fights. In our preferred specification (Panel A, column 4), a fight is associated

[18] Table A3 examines the sensitivity of results to alternative measures of gang scale. Specifically, I replicate the fixed-effects regressions from Table A2 using four alternative definitions. The benchmark measure is the number of locations in which the gang is present. The second measure is the (standardized) sum of the population living in those locations, capturing differences in population density. The third measure weights the number of locations by the estimated average consumer share for different drugs (cocaine, heroin, marijuana), accounting for heterogeneity in consumer distribution across areas; consumer numbers are estimated from population data and consumption propensities (please refer to Appendix D for additional details). The fourth measure weights the number of locations by the share of seized quantity. Results are consistent across all definitions.

[19] Are larger gangs less violent? Figure A9 addresses this question by comparing four metrics for small and large gangs. First, the proportion of fights involving large gangs is much lower: 85% of fights involve only small gangs. This partly reflects the smaller number of large gangs. Second, the annual probability that a gang becomes involved in a fight is higher for smaller gangs. Third, the number of fights per gang per year is higher for large gangs, likely because their broader geographic reach increases conflict opportunities. Fourth, fighting propensity–defined as the number of fights divided by the number of areas where the gang operates–is significantly lower for large gangs relative to their presence, consistent with larger gangs being less violent.

[20] Given that we only observe arrests at the city level, this is divided by 30–the number of districts–within Naples.

with a fivefold increase in gang homicides. The effect is even larger when at least one large gang is involved. In our preferred specification (columns 9–10), fights involving only small gangs are associated with a 437% increase in gang homicides, whereas those involving at least one large gang correspond to a 592% increase. The observed increase in violence is consistent with a more intensive deployment of resources during fights.

Districts in Naples are geographically small, and violence frequently spills across district boundaries regardless of gang presence, implying a SUTVA violation. Within Naples, during periods when a large gang is involved in a fight, gang homicides are 8.3% higher, on average, even in districts where the gang is not present. Outside Naples, this is 1.3%. For small gangs, spillovers are negligible: in Naples, districts without the fighting gang are associated with a 1.5% decline in homicides, while outside Naples the increase is only 0.2%. These means are reported in Table A5. Aggregating to the city level may mitigate this problem. For this reason, Tables A6 and A7 replicates Table A4 using city–year and gang–year data, respectively. Results are consistent.

Table A8 examines fighting spillovers in greater detail. Using gang–location–year data, we regress an indicator for whether a gang is involved in a fight in a given location-year on an indicator for whether the same gang is simultaneously engaged in another fight elsewhere. The probability of observing a fight in a given location increases significantly when the gang is also fighting elsewhere, even after including fixed effects (columns 1–4). Columns 5–9 turn to the number of gang homicides to examine violence spillovers. A fight in another location by the same gang is associated with a 43% increase in homicides in locations where the gang is active. Spillovers arise for both large and small gangs, though the effects are stronger for large gangs (36% vs 55%), despite the potential dilution from their broader presence across multiple areas. In column 9, the sample is restricted to gang–location–years without recorded fights. Even in this case, a fight elsewhere involving the same gang during the same year increases homicides locally, with larger effects when the fight involves a large gang. These results indicate that both fights and the violence they generate spill over across areas.

Violence appears costly for gangs, as reflected in subsequent increases in arrests, although the estimates are often imprecise. Using the same arrest specifications as with homicides (Table A4, Panel B, columns 4, 9, and 10), I find that a general fight is associated with a 6% increase in arrests; fights involving only small gangs with a 2% increase; and fights involving at least one large gang with a 11.57% increase. These estimates should be interpreted cautiously, since fights may simultaneously suppress illegal activities, biasing the measured effects downward. I return to this concern in Section 5.

Fact 2: Fights reduce the number of active gangs, especially when large gangs are involved

Approximately 50% of fights are followed by a reduction in the number of gangs active in the same location, either in the year of the fight or in the subsequent year. Figure A10 illustrates the percentage change in the number of active gangs in a location between the fighting period and the subsequent peaceful period. On average, a fight is associated with a 3% decrease in the number of active gangs during the following peaceful period. When distinguishing by the scale of the gangs involved, fights that include at least one large gang are associated with an 8% reduction in the number of active gangs, which is statistically different from zero at the 90% confidence level. Table A9 examines this relationship in more detail by estimating how the number of active gangs changes following a fight. Panel A includes all fights, Panel B restricts to fights involving only small gangs, and Panel C to those involving at least one large gang. Columns (1)-(3) present raw correlations; columns (4)-(6) add location and year fixed effects; and columns (7)-(9) further control for the lag of the dependent variable to account for pre-fight levels. Contemporaneous fights are associated with an increase in the number of active gangs, likely reflecting reverse causality. However, in the most complete specification (column 9), a fight is associated with a 2% reduction in the number of active gangs. This average effect masks substantial heterogeneity: fights involving at least one large gang are associated with an 8.5% decline in the number of active gangs in the affected location.

Table A10 complements this analysis by examining the probability that a given gang remains active in a location after a fight. Panel A considers all gangs, Panel B small gangs, and Panel C large gangs. Columns (1)-(3) report estimates without fixed effects (though controlling for the lagged dependent variable), while columns (4)-(6) include location, year, and gang fixed effects. Being involved in a fight is associated with a 12% increase in the probability of exiting the location in the following period. For large gangs the corresponding probability is only 1.5%.

Overall, fights are associated with a reduction in the number of active gangs in a location, potentially increasing the market share of the remaining groups. These estimates likely capture only the most extreme margin–complete exit from a location–while more moderate forms of territorial contraction (e.g., partial loss of control) cannot be observed in the data. Consequently, the effects may represent a lower bound. The magnitude of these reductions also varies systematically with gang scale, indicating that the fighting outcome probabilities may be heterogeneous.

To explore further this possibility, Table A11 summarizes the relationship between gang scale and fighting outcomes. Panel A compares the outcome probabilities of large and small gangs, estimated using an OLS regression of a fight outcome indicator (win, draw, lose) on a large-gang dummy variable. The results indicate that large gangs are significantly more likely to win fights, primarily because they are less likely to end in a draw. Panel B extends the analysis by replacing the binary large-gang indicator with a continuous measure of gang scale. A one-unit increase in gang scale is associated with a 3.6% increase in the probability of winning a fight (relative to a 10% baseline probability for gangs with a scale of one) and a 2.6% decrease in the probability of a draw. Panel E explores the relationship between gang size and the number of simultaneous conflicts. Larger gangs tend to be engaged in more contemporaneous fights, suggesting that once they enter a conflict, they are more likely to operate on multiple fronts. Conditional on participating in at least one fight during a given period, small gangs are involved in an average of 1.549 fights, while large gangs engage in 2.080, implying 0.521 additional fights for the latter. These spillover effects highlight a dynamic of conflict propagation, in line with the evidence presented in Fact 1. Panels C and D incorporate this feature by controlling for the total number of fights in which each gang is involved. The main findings remain robust: larger gangs continue to exhibit a higher probability of winning. I consider this empirical regularity as a primitive of the model, without attempting to endogenize it. Nonetheless, the evidence is consistent with the interpretation that larger gangs possess and deploy greater resources in combat, in line with the broader pattern of intensified violence documented in Fact 1.

Fact 3: Violence disrupts drug markets

Table A12 reports correlations between drug seizures and violence. Panel A presents regressions of the logarithm of the average quantity of drugs seized on the number of gang homicides. Panel B excludes observations above the 99th percentile to assess robustness to the exclusion of large seizures. In Panel C, the dependent variable is the share of a given drug–year seized in a city, and Panel D adds controls for the number of seizures. Panel E examines effects on the percentage of pure substance. Two considerations are important for interpreting these correlations. First, higher violence may attract greater police attention, increasing the number of seizures and biasing coefficients upward. To address this, the preferred specifications use the average rather than the total quantity seized and, in Panel D, explicitly control for the number of seizures. Second, violence may be endogenous to unobserved factors affecting drug markets. Both issues are addressed in the estimation strategy presented in Section 5.

Without fixed effects, higher homicide levels are positively correlated with the amount of drugs seized. This relationship disappears once the number of seizures is controlled for, suggesting that the initial correlation reflects increased enforcement activity. After conditioning on seizures, the coefficient becomes negative and statistically significant at conventional levels. This result is robust to the inclusion of city fixed effects (column 2) and city, drug, and year fixed effects (column 3) across all panels. In the preferred specification (Panel B, column 3), one additional gang homicide is associated with a 12% decrease in the average quantity of drugs seized. Results are quantitatively similar when using fights and their lag as explanatory variables (columns 4–6), though estimates are less precise. In the preferred specification, a fight is associated with a 4.5% decline in seized quantities in the contemporaneous period and a 13% decline in the following year. Interestingly, I also find an increase in the average purity of drugs during violent periods, consistent with the idea that higher quality is used to attract consumers, followed by a 10% decline after the conflict, in line with weaker competition (Panel E). However, quality results should be interpreted cautiously given the small sample.

Overall, the evidence indicates that violence disrupts drug markets. Periods with higher homicide or fighting activity are associated with lower quantities of drugs seized, consistent with a temporary reduction in trafficking during violent episodes.

3.4. Share vs margin trade-off

These features of the fighting technology suggest that gangs face a trade-off when deciding whether to engage in violence. On one hand, fighting may force rivals to abandon locations, and expand market share. On the other hand, they generate violence, which disrupts the market and attracts police attention. The relationship between market structure and violence depends on how changes in the first affect the incentives underlying this trade-off. Market structure can vary along two margins: (i) the number of gangs and (ii) their relative size distribution.

Consider a reduction in the number of active gangs, holding their relative sizes constant. Fewer gangs raise the returns from fighting, since more market share is at stake, but simultaneously reduce the opportunities for conflict. The net effect on violence is therefore ambiguous.

Now consider an increase in size asymmetry across gangs, holding the number of competitors constant. Larger gangs enjoy a higher probability of winning, increasing the expected benefits of violence. Smaller gangs, facing low chances of success, may avoid confrontation. At the same time, larger gangs recognize that conflicts generate more severe violence outbreaks and greater market disruption, raising the costs of fighting. To avoid regulatory attention, they may choose restraint, as in the self-regulation literature (Erfle & McMillan, 1990). Once again, the net effect of this change in concentration on the fighting trade-off is a priori ambiguous.

Therefore, the effect of changes in market structure on the use of violence is a priori ambiguous. It depends on the balance of several opposing forces and on key structural parameters governing the trade-off–such as the demand elasticity to violence, the intensity of police response, the probability of winning a fight, and the costs associated with increased police attention. To address this question, I develop a quantitative model of violent competition, estimate these parameters, and use counterfactual changes in market structure to study their implications for the use of violence.

4. A MODEL OF VIOLENT COMPETITION

This section develops a model of violent competition. Consumers choose where to buy drugs by weighing transport costs and local violence within a nested logit framework, following Goldberg, 1995. Gangs maximize expected profits by deciding whether to fight and whom to target. In doing so, they balance the gains from increased market share against the losses from reduced demand and heightened police attention. This trade-off parallels the predation and entry-deterrence literature, where firms deter rivals through capacity overinvestment (Spence, 1977; Dixit, 1980), predatory pricing (Milgrom & Roberts, 1982; Benoit, 1984), or other strategic actions (Schmalensee, 1978; Fudenberg & Tirole, 1986). We embed this logic in static entry models, building on Ciliberto and Tamer (2009) and Aguirregabiria and Ho (2012). The result is a model of directed competition in which firms may undertake costly actions to expel rivals and capture market share.

4.1. The basic setup

Consider a finite set of locations $y = 1, 2, \ldots, Y$, gangs $j = 1, 2, \ldots, J$, and consumers $i = 1, 2, \ldots, I$, of drugs $v = 1, 2, \ldots, V$ over periods $t = 1, 2, \ldots, T$. The model is static, justified by the need to preserve tractability in computing an equilibrium with many gangs and locations. Let $\mathbf{Y}_t$ denote the location matrix, where $y_{jt} = 1$ if gang j is present in location y at t. On the extensive margin, we assume that a gang active in a location can produce all drugs.[21] This assumption substantially simplifies the state space, requiring only one location matrix per period instead of V matrices.

4.2. Demand

Utility Consumer i, residing in location y_i, chooses at time t whether to purchase drug v from gang j in location y, or to select the outside option of no purchase ($y = 0$). The conditional indirect utility from buying drug v from gang j in y is

$$u_{ijyvt} = \alpha_v p_{vt} + \mu_v V_{yt} + \xi_y + \xi_{vt} + \xi_{yvt} + d_v g_{iy} + \epsilon_{ijvt} + (1-\rho)\epsilon_{iyvt} \tag{1}$$

where p_{vt} is the average price of drug v at time t, and V_{yt} denotes violence in location y at time t. Therefore, we allow consumer to value violence in the locations, in line with Fact 3 and the insights of Levitt and Venkatesh (2000).[22] The terms $\xi_y, \xi_{vt}, \xi_{yvt}$ represent unobserved (to the econometrician) location, product, and location–product–time preferences, respectively, observed by both suppliers and consumers. g_{iy} is the distance between consumer i and location y. The terms ϵ_{ijvt} and ϵ_{iyvt} are consumer-specific taste shocks for drug v from supplier j and from location y, respectively. Both are assumed to be identically type-I extreme value distributed.

Define nests as $g = 0, 1, \ldots, Y$ for each drug v, where $g = 0$ denotes the outside option (with $g_{i0} = 0$ for all i), and $g = 1, \ldots, Y$ correspond to locations within the province of Naples. The nesting parameter ρ governs substitution patterns across locations. I impose $\rho = 0$, implying perfect substitutability across nests. In this framework, consumers are not permitted to switch across drug types.

Consumer problem Let $d^{v,t}_{(y,j)}$ denote the indicator of consumer choice, with $d^{v,t}_{(y,j)} = 1$ if consumer i of drug v selects bundle (y, j) at t and 0 otherwise. The consumer's problem can then be written as:

$$\max_{d^{v,t}_{(y,j)}} \sum_{(y,j)} d^{v,t}_{(y,j)} u_{ijyvt} \quad \forall v, t$$
$$\text{s.t.} \sum_{(y,j)} d^{v,t}_{(y,j)} = 1$$

[21] This assumption is consistent with the data. Seizures of cocaine almost always coincide with seizures of heroin and marijuana (Figures A11 and A12): no spatial specialization in drug trafficking. In fact, all drugs are sold in the same areas. In principle, if a particular drug is not profitable in a location, gangs can always adjust on the intensive margin by setting its quantity to zero.

[22] It is important to mention here that drug consumption is not a major criminal offense in Italy. Under Italian law (Presidential Decree No. 309/1990, the Consolidated Law on Drugs), drug use is not a criminal offense. Possession for personal use is subject only to administrative sanctions (e.g., suspension of driving licenses or firearms permits), and for first-time offenders typically results only in a warning and a request to refrain from use. See https://canestrinilex.com/en/readings/drug-offences-in-italy-detention-and-dealing.

In this framework, an individual is fully characterized by product-specific shocks $(\epsilon_{ijvt}, \epsilon_{iyvt})$ and by location y_i. This structure implicitly defines the set of individual attributes that generate demand for bundle (j, y) from y_i. For ease of exposition, the choice problem may be viewed as a two-step process: first, an across-nest decision, where the consumer selects the location in which to buy; second, a within-nest decision, where the consumer selects the supplying gang.

Within nests The within-nest decision is straightforward. Conditional on an idiosyncratic taste shock, each consumer selects the gang that maximizes utility. Since consumers differ only by origin location, and we focus on those who have already chosen their optimal area, the share of consumers purchasing from a given gang is independent of i. Moreover, because the model abstracts from price and product differentiation–due to data limitations–this share reduces to the reciprocal of the number of active gangs in the location.[23] Appendix A.1 derives the conditional demand formally.

Across nests In the across-nest decision, geographically dispersed consumers weigh unobserved preferences, transportation costs, and the disutility from violence. Exploiting the properties of the type-I extreme value distributed consumer-specific shock ϵ_{iyv}, I derive the across-nest demand function (please refer to Appendix A.1 for details).

Consumer sorting across locations plays a central role in demand estimation, given the structure of the data. Specifically, I use variation across, rather than within, nests to identify demand parameters. To bring the model to the data, I aggregate the unconditional demand for nest y as a weighted sum of conditional demands, take logarithms of the resulting shares, and apply standard log properties. This yields a log-linear specification of the form:

$$log(s_{yvt}) = \mu_v V_{yt} + f_y + f_{vt} + \xi_{yvt} \tag{2}$$

which is the reduced-form demand equation used in estimation in Section 5.

Unconditional demand Combining the conditional within-nest demand with the across-nest de-

[23] Detailed price data at the gang–location level are not observed. For this reason, I choose not to model within-location price competition. As a result, consumers are assumed to select gangs at random once they have chosen a location. This also implies that we cannot capture changes in market concentration arising from unequal gang-specific market shares, as already mentioned in Section 3.

mand yields the unconditional demand for bundle (y, j) of drug v at time t.

4.3. Supply

Gang problem At the beginning of each period t, each gang $j = 1, 2, \ldots, J$ observes the initial network matrix $\mathbf{Y}_t$, and hence the associated vector of gang scales $\mathbf{m}_t$, and simultaneously solves a two–stage game:

1. In the first stage, the gang decides whether to fight (0/1), where to fight (y), and against whom (o). Let $f_{joyt} = 1$ denote the decision to fight gang o in location y. Stacking the $Y \times J$ possible decisions of gang j at t yields the fighting matrix $\mathbf{F}_{jt}$. The collection of these matrices across gangs, together with the realization of fight outcomes, determines the updated network matrix $\mathbf{Y}_t \to \mathbf{Y}'_t$.
2. In the second stage, conditional on the realized network matrix, demand, violence, and police attention, profits are realized.

It is important to note that the model does not explicitly account for coalition formation, for two reasons. First, the decision not to fight can itself be interpreted as a form of (implicit) coalition. Second, explicit alliances in this context tend to be highly unstable, as extensively documented in intelligence reports (DIA, 2015 to DIA, 2022). Nonetheless, this should be regarded as a limitation of the model.

Dyadic representation It is useful to introduce an alternative formulation of the game. Define a *dyad* as a triple $d_t = (j, o, y)$. A fight occurs in dyad d_t if $f_{dt} = f_{joyt} f_{ojyt}$. The entire game can be represented as a collection of interconnected dyads, with gangs making strategic choices within each of it. The scale of dyad d at time t is defined as $m_{d_t} = \max m_{jt}, m_{ot}$. With a slight abuse of notation, I write $y \in d_t$ to indicate that location y belongs to dyad d at time t, and apply the same convention to gangs.

Some key assumptions To reduce the choice space and accommodate salient features of the data, I impose four key assumptions:

I. Gangs are restricted to fight only in areas where they are active during the period (accounting for 86% of fights in the data);

II. Fights may end with either a definitive outcome (one gang wins and the other loses) or a draw. This allows me to model fights lasting more than one period as repeated draws;

III. I do not impose restrictions on the number of fights a gang may undertake within a period. I assume, however, that the probability of winning any given fight is independent of the number of concurrent fights. This permits me to treat multi-party fights (approximately 40% of the sample) as collections of independent bilateral fights.[24]

IV. In principle, each gang faces a *joint* simultaneous decision problem across all dyads in which it is active. For instance, a feasible strategy could be to fight in only one dyad, irrespective of which one. I assume that fighting decisions are taken *separately* within each dyad. Formally, each dyad constitutes an independent choice problem for the gang, although payoffs remain interdependent across dyads through violence spillovers and police attention. This assumption rules out strategies involving coordinated choices across dyads, while preserving interdependence through equilibrium outcomes.[25] Consistent with this interpretation, I assume that the fixed cost of fighting is dyad-specific, privately observed within that dyad, and not known even to the same gang when making decisions elsewhere. This assumption significantly reduces the dimensionality of the gangs' choice space. Intuitively, it can be rationalized by limited coordination capacity: gangs may be unable to perfectly synchronize contemporaneous fighting decisions across dyads, for example due to "heat-of-the-moment" shocks, and instead form expectations about outcomes in other dyads when choosing optimally in each one.

These assumptions will be formally discussed in the following paragraphs. Given the similarity between this fighting game and a classical static entry model, following Reiss and Bresnahan (1990), Berry (1992), and Mazzeo (2002), I do not consider mixed strategy equilibria.

Choice set Let $\mathcal{D}_{jt} \equiv \{d_t : j \in d_t\}$. In principle, the choice set of gang j at time t would be:

$$\mathcal{F}_{jt} = \prod_{d_t \in \mathcal{D}_{jt}} \{0, 1\} = \{0, 1\}^{|\mathcal{D}_{jt}|}$$

[24] This assumption is partially supported by the evidence in Table A11, which shows that fight outcomes are not systematically correlated with the number of simultaneous fights involving a gang.

[25] A similar local-decision structure is employed in Aguirregabiria and Ho (2012), where airlines are modeled as delegating entry decisions to market-specific managers. An alternative assumption often used in the entry literature to reduce dimensionality of the choice set is that firms may enter only one location per period, as in Aguirregabiria and Vicentini (2016). This restriction is not suitable for this model because it would implicitly impose a lower fighting propensity on larger gangs.

Under Assumption IV, however, gangs do not optimize over this joint set. Instead, each dyad generates a separate binary decision problem with choice set $\mathcal{F}_{jd_t} = \{0,1\},\ d_t \in \mathcal{D}_{jt}$. Hence, the choice set of gang j in period t would consist of the collection of these per-dyad sets:

$$\mathcal{F}_{jt} = \{\mathcal{F}_{jd_t}\}_{d_t \in \mathcal{D}_{jt}} = \{\{0,1\},\ldots,\{0,1\}\}$$

Moreover, under Assumption I, I further restrict the choice set by restricting the number of potential choice-dyads. Define $\mathcal{D}'_{jt} \equiv \{d_t : j \in d_t \text{ and } y \in \mathcal{Y}_{jt} \text{ and } y \in \mathcal{Y}_{ot}\}$. To conclude, note that the information structure separates decisions across dyads, even when they involve the same gang. It is therefore natural to define the choice set at the *gang–dyad* level rather than at the gang level. In this formulation, each gang–dyad is treated as an independent player, though all gang–dyads belonging to the same gang maximize the same objective function. Denote by j_{d_t} the player corresponding to gang j in dyad d_t. The choice set of gang–dyad j_{d_t} is simply $\mathcal{F}_{j_{d_t}t} = \{0,1\}$.

Production Gangs produce drugs using pure drug r and cutting substances s according to a Leontief production function:

$$q_{jyvt} = \min\{\frac{s_{jyvt}}{1-x_v}, \frac{r_{jyvt}}{x_v}\} \tag{3}$$

x_v is the inverse of the average percentage of pure drug in the dose, and it estimated from the data. This production function is very intuitive: to produce $q_{jyvt} = 1$ gram of drug, the gang needs x_v grams of pure substance and $1 - x_v$ grams of cutting substance.

Police Police is a function:

$$O_{yt} = \bar{o}_y + f(\lambda V_{ty}) \tag{4}$$

where $\bar{o}_y$ is the average police attention in y, V_{ty} is the amount of violence in location y at time t. Therefore, police in this model reacts to violence by allocating attention.

Violence Violence is determined according to the following law of motion:

$$V_{yt} = \sum_{\{d\,:\,y \in d_t\}} \theta_m 1\{m_{d_t} = m\} f_{d_t} + \sum_{\{j\,:\,y \in \mathcal{Y}_{jt}\}} \sum_{\{d_t\,:\,j \in d_t\ y \notin d_t\}} \psi_m 1\{m_{dt} = m\} f_{d_t} \tag{5}$$

In words, violence is a step function that increases heterogeneously with respect to both: (i) fights that occur directly at location y, with intensity parameter θ_m; and (ii) fights involving gangs active in y that take place elsewhere, with spillover parameter ψ_m. This formulation captures the violence dynamics described in Fact 2.

Second stage Profits in the second stage are:

$$\Pi(\mathbf{Y}_t, \mathbf{V}_t, \mathbf{F}_{jt}) = \sum_{\{y \in \mathcal{Y}_{jvt}\}} \{\sum_{\{v\}} [p_{vt} - [k_{vt} + (c_{vt} - k_{vt})x_v]] q_{jyvt}(\mathbf{V}_t) - \delta O_{yt}(\mathbf{V}_t)\} - \boldsymbol{\gamma} \mathbf{F}_{jt} \quad (6)$$

where δ captures costs from police enforcement, rationalized by arrests, drug seizures, or other enforcement activities. This cost enters as a fixed cost, independent of quantities, and should be interpreted as the cost of operating in a given area. The term $\gamma_{joyt} \sim^{iid} N(\bar{\gamma}, \sigma_\gamma)$ is a random fixed cost of fighting, drawn at the beginning of each period and privately observed by both gangs in the dyad during the first stage. These realizations become observable in the second stage. The random cost captures both the non-rational or stochastic component of fighting and not modeled costs (e.g., the cost of lives). In the context of a standard entry game, this term parallels an heterogeneous entry costs. Assumptions I and IV can be formally written as:

Assumption I and IV. *The choice set of gang j at time t is $\mathcal{F}_{jt} = \{\mathcal{F}_{jd_t}\}_{d_t \in \mathcal{D}'_{jt}}$ where $\mathcal{D}'_{jt} \equiv \{d_t : j \in d_t$ and $y \in \mathcal{Y}_{jt}$ and $y \in \mathcal{Y}_{ot}\}$. Relatedly, the shocks γ_{jyot} are private information of both gangs j and o in the dyad $d_t = (j, o, y)$. These shocks are unknown to the same gangs in other dyads.*

Assumption V. *γ_{jyot} are independently and identically $N(\bar{\gamma}, \sigma_\gamma)$ distributed for each gang-dyad j_{d_t}.*

Assumption V is important, since it rules out dependence between fighting shocks both within and across dyads.

Fighting mechanism Building on Donohue and Levitt (1998), each gang-dyad j_{d_t} is endowed with a fighting ability ϕ_{jt}, composed of an observable component ϕ_j, common knowledge among all gangs, and an unobservable shock ϵ_{jyot}. The latter captures randomness in fight outcomes and is assumed to be independently and identically ETV1 distributed with location 0 and scale σ_ϵ. The parameter σ_ϵ thus measures the degree of unpredictability in fighting outcomes. Consistent with Fact 2, we allow fighting ability–and therefore the probability of winning–to vary across gangs, reflecting differences in gang scale. Since the observable component of fighting ability is common knowledge, along with all other profit components, the game is one of complete information. A natural extension would be to allow for incomplete information by treating fighting ability as private information.

Suppose a fight between a gang j and o at time t in y. Fighting mechanism is a step function:

$$\begin{cases} \phi_j + \epsilon_{jyot} - \phi_o - \epsilon_{oyjt} > \Phi & j \text{ wins} \\ \Phi \geq \phi_j + \epsilon_{jyot} - \phi_o - \epsilon_{oyjt} > -\Phi & \text{draw} \\ -\Phi \geq \phi_j + \epsilon_{jyot} - \phi_o - \epsilon_{oyjt} & j \text{ loses} \end{cases} \tag{7}$$

where Φ is the step parameter. Formally, assumptions II and III can be written as:

Assumption II. *Fights may end with a draw: $|\Phi|>0$.*

Assumption III. *For any gang j in any dyad $d_t = (j, o, y)$, the fighting shock ϵ_{jyot} is independently and identically ETV1 distributed with location 0 and scale σ_ϵ.*

The game The timing of the game is as follows. At the beginning of each period, each gang–dyad j_{d_t}, for $d_t = (j, o, y)$, observes the current network matrix $\mathbf{Y}_t$ and its idiosyncratic taste shock γ_{jyot}. Conditional on this information, the two gangs within the dyad sequentially decide whether to fight, with the order of moves randomly determined by nature. The sequential structure is central. In a simultaneous-move game, the unique Nash equilibrium of these fighting games is typically mutual fighting: given strategic uncertainty about the opponent's type and action, fighting becomes dominant strategy, leading both players to this outcome (Baliga & Sjöström, 2004, 2012). Such a formulation rules out cooperative or deterrence-based outcomes. Sequential play mitigates this problem by allowing signaling and credible threats.[26] Because of the sequential structure, the relevant solution concept is Subgame Perfect Nash Equilibrium (SPNE). The game is illustrated in Figure A14. A pure strategy for gang–dyad j_{d_t} is a mapping from states and shocks into a binary action:

$$\sigma_{j_{d_t}}(\mathbf{Y}_t, \gamma_{joyt}) \in \{0,1\} : \{0,1\}^{Y \times J} \times \mathbb{R} \longrightarrow \{0,1\}$$

Once strategies are chosen, fighting outcomes are realized, the network updates to $\mathbf{Y}'_t$, and payoffs are determined in the second stage.

Consider the set of all possible states of the world in each dyad d_t

[26] An alternative modeling device would be to impose a fixed cost of fighting whenever a player threatens to fight, regardless of whether a fight actually occurs. We view this as a second-best solution.

$$\omega \in \Omega = \{\text{WNF,WF,P,D,NF,L}\}$$

which indicate, respectively, win without fighting, win with fighting, peace, draw, non fight, and loose. Define an implicit profit function:

$$\Pi(\mathbf{Y}_t'(\boldsymbol{\omega}), \mathbf{F}_t(f_{j_{d_t}}, f_{-j_{d_t}}, \mathbf{F}_{j-d_t}, \mathbf{F}_{-j})) = \pi : \{0,1\}^{Y\times J} \times \{0,1\}^{Y\times J\times J} \longrightarrow \mathbb{R}$$

and a fighting outcome function:

$$P(\boldsymbol{\phi}, f_{j_{d_t}}, f_{-j_{d_t}}, \omega) = p : \mathbb{R}^J \times \{0,1\} \times \{0,1\} \times \Omega \longrightarrow [0,1]$$

which, for each vector observed fighting abilities, fighting decisions, and possible state reports the probability of that state happening. Then, gang-dyad j_{d_t} solves:

$$V(\mathbf{Y}, f_{j_{d_t}}, f_{-j_{d_t}}, \mathbf{F}_{j-d_t}, \mathbf{F}_{-j}) = \max_{f_{j d_t}} \sum_{\{\boldsymbol{\omega}\}} P(\boldsymbol{\phi}, f_{j_{d_t}}, f_{-j_{d_t}}, \omega)\mathbb{E}\Pi(\mathbf{Y}_t'(\boldsymbol{\omega}), \mathbf{F}_t(f_{j_{d_t}}, f_{-j_{d_t}}, \mathbf{F}_{j-d_t}, \mathbf{F}_{-j})) \tag{8}$$

A simpler game Consider first the problem of a single dyad (j, o, y) in isolation. To simplify notation, let the profits in each state of the world be:

$$\pi_{WNF}, \pi_{WF}, \pi_P, \pi_D, \pi_{NF}, \pi_L$$

The game begins with nature assigning a random opportunity to fight to one of the gangs. Suppose gang j is chosen. Gang j then decides whether to fight ($F = 1$) or not ($NF = 0$). If j chooses NF, both gangs obtain the peace payoff π_P. If instead j chooses F, then it is o's turn to respond. If o declines to fight, it exits the area. In this case, j receives the win–non-fight payoff π_{WNF}, while o obtains the non-fight payoff π_{NF}. If, on the other hand, o chooses to respond, both gangs obtain the expected fighting payoff. For j, this is given by $\mathbb{E}\pi_F = P(\phi_j + \epsilon_{jyot} - \phi_o - \epsilon_{oyjt} > \Phi)\pi_{WF} + P(\phi_j + \epsilon_{jyot} - \phi_o - \epsilon_{oyjt} < -\Phi)\pi_L + (1 - P(\phi_j + \epsilon_{jyot} - \phi_o - \epsilon_{oyjt} > \Phi) - P(\phi_j + \epsilon_{jyot} - \phi_o - \epsilon_{oyjt} < -\Phi))\pi_D$ for j, and similar by changing the probabilities and relevant profits for o.

In Appendix A.2 I present the SPNEs for the game with uncertainty.[27] Nine SPNEs emerge. These are summarized in Figure 1. In many cases, the actual equilibrium depends on nature's choice, creating

[27] Appendix A.2 also presents the SPNEs for the case with no uncertainty, $\sigma_\epsilon = 0$. In that case, six possible equilibria arise depending on parameter values and the distribution of second-stage profits. These six equilibria map into three scenarios: (i) j exits, (ii) o exits, (iii) multiple equilibria. Specifically, j exits whenever o is expected to win the fight, or when they draw but o can credibly sustain the threat of fighting. Symmetrically, o exits under the opposite conditions. Finally, multiple equilibria occur either when they draw and neither gang can credibly threaten to fight (in which case the first mover wins the area), or when both can credibly threaten (in which case the first mover exits). Importantly, across all scenarios the unique predicted outcome is no fighting, in line with the conflict literature under certainty.

multiplicity (highlighted in bold green). To resolve this multiplicity, I adopt a natural selection rule: focus on the equilibrium that would emerge as an absorbing state if the game were (non-strategically) repeated.[28] As the number of repetitions increases, the probability of nature's move ceases to matter. This procedure yields five scenarios: (i) j exits whenever expected profits from fighting are lower than from not fighting and o can credibly threaten to fight; (ii) o exits under the symmetric condition; (iii) multiple equilibria appear when fighting yields lower profits than not fighting for both gangs, so neither can credibly threaten–here, the first to move secures the area; (iv) nothing happens when expected profits from fighting are higher than from not fighting for both gangs, but never exceed peace profits–in this case, both can credibly threaten, so neither chooses to fight; (v) a fight occurs when both gangs are willing to fight conditional on the rival fighting, and at least one has expected profits from fighting strictly greater than from peace. In this last case, both gangs can credibly threaten to fight, and since fighting is strictly optimal for at least one of them, a fight eventually occurs. The conditional choice probabilities predicted by the model are summarized in Appendix A.2.

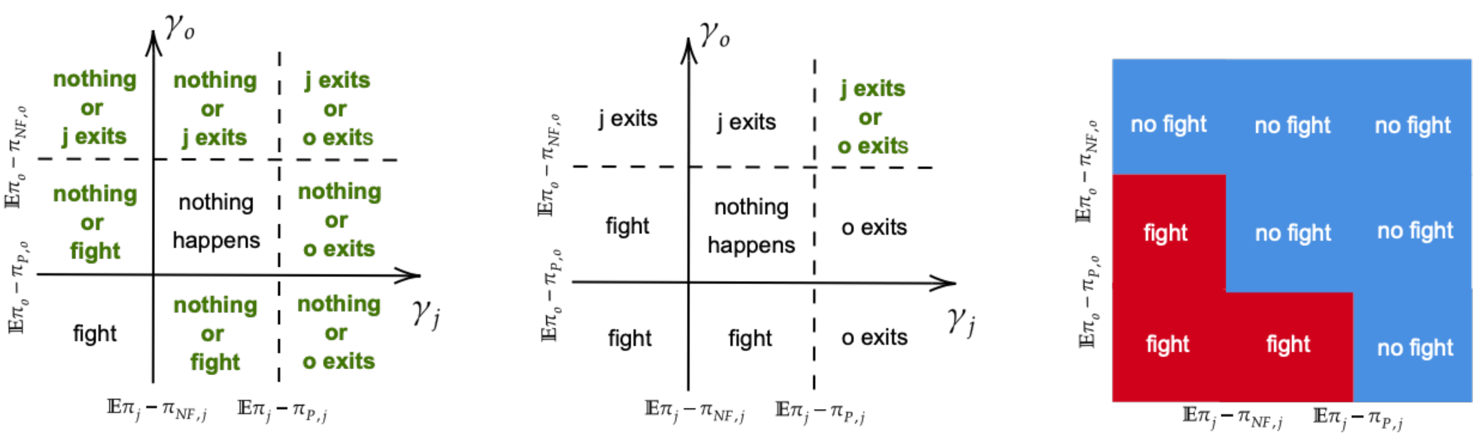


Figure 1: Fighting outcomes

Although this simple game admits multiple equilibria, our interest is in the incidence of fighting, which is uniquely predicted within each dyad.[29] Fighting probabilities are summarized in Appendix A.2.

The extended game The extended game is a collection of single-dyad games, whose payoffs are linked

[28] For example, in the bottom-right case, if nature selects gang j, then o exits the market; if instead nature selects o, nothing happens. Repetition ensures that eventually j is selected, forcing o to exit. The absorbing state is thus uniquely determined.

[29] This is equivalent to the argument presented in Berry (1992).

through aggregate violence, the corresponding allocation of police attention, and fight taste shocks.[30] Conditional on the level of violence–and thus on police attention–the only dependence across dyads arises from taste shocks in the profits of the same gang. Since these shocks are additively separable, they do not affect the marginal decision of whether to fight in other dyads. This insight provides a useful simplification. I can reformulate the problem by treating each dyad as a player facing a game against nature: given the prevailing level of violence, the dyad's problem uniquely determines the probability that fighting is the optimal action for its member gangs. It is worth emphasizing that this reformulation preserves the strategic interdependence of fighting choices, as payoffs are interdependent across dyads. Hence, the overall multi-dyad game continues to admit multiple equilibria. Nonetheless, the reformulation guarantees that each dyad's best response to aggregate violence is uniquely determined, thereby considerably simplifying the characterization of equilibrium behavior and the subsequent estimation of the model.

4.4. Equilibrium

Consider the following three vectors:

- Parameters: $[p_{vt}, c_{vt}, k_{vt}, \alpha_v, d_v, \mu_v, \xi_y, \xi_{vt}, \xi_{yvt}, x_v, \lambda, \bar{o}_y, \theta_m, \psi_m, \gamma, \sigma_\gamma, \delta, \phi, \sigma_\epsilon, \Phi]$
- State: Y_t
- Choice: $\mathbf{F}_t$

An equilibrium is a matrix of fight choices $\mathbf{F}_t$ such that, given parameters and state, this is a Sub-game Perfect Nash Equilibrium of the fighting game (8) for all gang-dyads.

5. IDENTIFICATION AND ESTIMATION

Data Table A13 summarizes the available data, their role in the model, and their sources. The data set includes network matrices (both ex-ante and realized, with the former reconstructed by reversing fight sequences), drug seizures, prices, and fighting decisions and outcomes. However, drug quantities and police attention are not directly observed. For these two variables, we rely on proxy measures that are inherently subject to selection. In the following paragraphs I discuss this selection.

[30] Although we model decisions within each dyad as independent, crucially, their payoffs are interdependent. The violence response and its spillovers in other locations play a "reduced form" role of fight interdependence.

Drug quantities Drug quantities are unobserved ($q_{yvt} \rightarrow q^*_{yvt}$). I model the selection process through the following equation:

$$Q_{vyt} = s(q^*_{vyt}, O^*_{yt}, \epsilon_{yvt}) = q^*_{vyt} s(O^*_{yt}, \epsilon_{yvt}),$$

where O^*_{yt} denotes police attention and ϵ_{yvt} captures additional factors affecting seizure rates, potentially correlated with demand and other market determinants. In other words, here I am assuming *separability* between true quantity and selection.

Police attention Police attention is unobserved ($O_{yt} \rightarrow O^*_{yt}$). In the data, I observe several potential measures, namely arrests for different categories of crimes. Let I_{yt} denote the incidence of illegal activity i in area y at time t. Arrests for activity i, A^i_{yt}, can be expressed as a function of both the incidence of crime and police attention:

$$A^i_{yt} = a(I_{yt}, O^*_{yt}).$$

This equation highlights the importance of selecting an appropriate measure for police attention. This is a measure where the incidence of the underlying activity is unaffected by violence, $\frac{\Delta \hat{I} yt}{\Delta V_{yt}} \approx 0$. In that case, changes in arrests for that activity capture changes in police attention: $\frac{\Delta A^{\hat{i}}_{yt}}{\Delta V_{yt}} = \frac{\Delta O^*_{yt}}{\Delta V_{yt}}$. I specify police attention as a linear function of violence: $O^*_{yt} = \bar{o}_y + \lambda V_{yt} + u_{yt}$.

Model simplification I impose a set of simplifying assumptions to bring the model to the data. First, I restrict the number of demand parameters by assuming homogeneous preferences across drugs, α, d, μ. Second, I reduce heterogeneity across gangs by distinguishing only two types–small and large–consistent with the descriptive evidence. Accordingly, the model includes two parameters for fighting ability (ϕ_s, ϕ_l), fight–violence elasticity (θ_s, θ_l), and violence spillovers (ψ_s, ψ_l). We normalize $\phi_s = 1$, so that ϕ_l is interpreted in relative terms.

Overview A subset of parameters is directly calibrated from the data, $p_{vt}, c_{vt}, k_{vt}, x_v$. The estimation proceeds in five steps. First, the demand parameters $\alpha, d, \mu, \xi_y, \xi_{vt}$ are identified from the correlation between seized drug quantities and variation in prices, distance, and violence, controlling for selection. This amounts to estimating a fixed-effects regression–exploiting the linear structure of unconditional demand across areas (Equation 2)–controlling for the share of seizures to address selection and an IV strategy to account for the endogeneity of violence. Second, conditional on observed fighting choices, fighting parameters $\psi, \phi_j, \Phi, \sigma_\epsilon$ are estimated by simulated method of moments, using data on fight outcomes. Third, police reaction parameters $\lambda, \bar{o}_y$ are identified from the correlation between arrests and gang homicides, controlling for fixed effects. Fourth, violence generation parameters $\theta_s, \theta_l, \psi_s, \psi_l$ are estimated from changes in gang homicides around the outbreak of fights. Fifth, the remaining parameters governing fighting decisions, $\gamma, \delta, \sigma_\gamma$, are estimated via a nested fixed point algorithm. In the following paragraphs, we informally discuss identification and provide estimation details for each step separately.

Preliminary step As a preliminary step, a subset of parameters is directly calibrated from the data: $p_{vt}, c_{vt}, k_{vt}, x_v$. Average retail and wholesale prices are taken from the official reports submitted by enforcement authorities to the national antidrug agency. The price of cutting substances is fixed at 0.01 euros per gram. Drug quality is set equal to the observed average for cocaine and heroin, and normalized to one for marijuana.

Step I: Demand parameters The identification of the consumer disutility from violence μ is subtle. A natural approach would be to estimate μ from changes in seized drug quantities associated with violence.[31] However, this naive approach is likely to produce biased estimates due to selection.

Changes in drug seizures following violence reflect two components: (i) the direct effect of violence on demand (quantity), and (ii) the change in selection arising from changes in seizure rates. Without separating these effects, the parameter μ is not identified. Identification requires controlling for the variation in seizure rates induced by police attention. I implement this using two alternative methods. First, I control directly for the share of seizures, building on the separability assumption. Second, I estimate a two-step control function: in the first stage, I regress the number of gang homicides on the share of arrests by the local police, which proxies for police attention; in the second stage, I estimate the log-linear demand equation, including the predicted residuals to control for the selection effect.[32]

[31] An alternative approach is to examine how the share of seized quantities varies with the outbreak of fights. Figure A13 plots the average log seized share, by drug and year, against years to the first fight, controlling for city fixed effects. The decline in seized share is visible in the mean and becomes more pronounced in fights involving at least one large gang, consistent with these fights generating greater violence.

[32] The credibility of this approach hinges on the validity of the police attention measure. The measure must be relevant for

Results are presented in Table A14 (columns 1–3). Across specifications, the coefficient on gang homicides is negative and statistically significant. Controlling for seizure shares (column 2) or fitted selection (column 3) increases the magnitude of the coefficient, consistent with downward bias in the naive estimates. In columns (4) and (5), I instrument for gang homicides with the number of gang members released from prison. The instrument is motivated by the idea that recently released members have incentives to avoid conflict. Prison releases thus shift gangs' propensity to engage in violence but should not directly affect drug demand. This parallels the idea of supply-side shifters, such as input prices, used as instruments in demand estimation strategies: they affect equilibrium prices but are excluded from consumers' utility. First-stage coefficients are negative and sizable, with KP statistics of approximately 70, well above conventional thresholds. The second-stage coefficients remain negative, similar or slightly larger in magnitude than the OLS estimates, and statistically significant at conventional levels. In column (5) I additionally control for the total number of homicides, and results seem to be driven by gang homicides specifically.

Unobserved preference parameters, ξ_{vt} and ξ_y, are recovered from the estimated fixed effects. Price elasticity is estimated via an IV regression of the drug–year fixed effect on retail prices, using as instrument the wholesale prices. The first-stage coefficient is positive, with a large KP statistic. The estimated price elasticity is small in magnitude. Finally, parameter d is recovered from the estimated fixed effect $\hat{f}_y$ (Panel C of Table A14).

Panel D examines the robustness of the estimates to alternative functional forms. It reports estimated coefficients and the corresponding elasticities for four specifications of the benchmark model: (i) the logarithm of the share plus one–used because many observations record no drug seizures; (ii) the logarithm of the share without the plus one; (iii) the logarithm of the share plus one regressed on the logarithm of homicides plus one; and (iv) the logarithm of the share without the plus one regressed on the logarithm of homicides plus one.[33] Coefficients are qualitatively similar across these specifications, but their magnitude varies substantially, due to changes in the sample composition.

seizures but excluded from the demand equation. For this reason, I use the standardized number of arrests by the local police. Unlike arrests for drug-related crimes, which may directly affect the drug market, local police arrests are relevant for seizures but unlikely to shift demand. As discussed in Section 2, this force focuses primarily on traffic control and urban enforcement, and drug use is not classified as a major crime in Italy. The exclusion restriction is therefore more plausible in this setting.

[33] The plus one in the logarithm of gang homicides is always retained because a large share of observations record zero homicides.

Step II: Fighting parameters The fighting parameters, ϕ_l, Φ, σ_ϵ, are estimated by simulated method of moments using data on fight outcomes. The estimation matches three key moments: (i) the higher probability of winning when a large gang is involved, (ii) the average probability that a fight ends in a win, and (iii) the variance in fight outcomes. The first moment identifies the relative fighting advantage of large gangs, ϕ_l. Conditional on the scale of the gangs engaged in the conflict, the second moment identifies the overall scale of the fighting process, Φ. Finally, conditional on scale and step, higher dispersion in fight outcomes reflects greater unpredictability in the mechanism; thus, the variance in winning probabilities identifies σ_ϵ.

Step III: Police reaction parameters The police reaction parameters, $\bar{o}$, λ, are estimated from the relationship between gang homicides and arrests. The identification of these parameters poses a challenging measurement problem, on top of the standard reverse causality concerns. Violence is correlated with both police activity and the incidence of crimes. Failing to account for this measurement problem would probably bias our estimates. Table A15 reports fixed-effects regressions of gang homicides on the share of arrests across different crime categories and enforcement agencies: arrests by local police, arrests recorded by the DIA, arrests for conspiracy, and arrests for drug-related offenses. The preferred measure of police attention is arrests for conspiracy: a log-run crime unlikely to respond strongly and immediately to violence. Columns (1)–(4) present estimates with city fixed effects, and columns (5)–(8) add year fixed effects.[34] To capture the dynamic nature of arrests, both current and lagged numbers of gang homicides are included.[35] The use of lagged homicides also helps mitigate reverse causality concerns. Panel A reports estimates using the number of arrests, Panel B uses the share of arrests, Panel C excludes Naples as a robustness check, Panel D replaces the homicide count with a dummy variable, and Panel E combines the dummy specification with the exclusion of Naples. The parameter λ is thus estimated as the total change, across the two periods, in the share of conspiracy arrests associated with changes in gang violence. Panel B shows an overall positive correlation–combining current and lagged coefficient–between violence and police attention for the first three arrest measures: local police, DIA, and conspiracy. By contrast, the coefficient for drug-related arrests is small and negative, consistent with earlier evidence on drug seizures. The average police attention parameter $\bar{o}$ is recovered from the estimated city fixed effects. District aver-

[34] Because the data are administrative, arrests are recorded at the city level rather than at the neighborhood level.

[35] See Tables A6 and A7 for the evolution of arrests during and after fights.

ages are obtained by dividing the citywide estimate by the number of districts in Naples, resulting in the same average across the city.

Step IV: Violence parameters The parameters governing violence generation, $\theta_s, \theta_l, \psi_s, \psi_l$, are identified from changes in gang homicides at the outbreak of fights. Estimation is conducted at the city level rather than the location level, in light of the geographical spillovers documented in Section 3.3.[36] Results are reported in Table A16, Panel A. The estimates are similar when excluding Naples from the sample (panel B). Columns (4) to (6) restrict the control group to location which have experienced at least one fight during the whole sample period.

Spillover parameters, ψ_s, ψ_l, are estimated using location-level variation, since the object of interest is precisely to measure spillovers. I regress the number of gang homicides in locations where a gang is active on an indicator for whether that gang engages in a fight elsewhere during the same period. Results are presented in Table A16, Panel B. In column (3), I restrict the control group to locations without fights in the period, thereby isolating spillovers from external rather than local fighting activity. All specifications include gang, location, and year fixed effects.

Table A17 replicates the results of Table A16, Panel A, using both location-level and city-level variation and including the full distribution of gang scale rather than the dichotomous small–large classification. Because few fights involve gangs with scale greater than five, columns (3) and (4) winsorize gang scale at five. Both panels show an increase in violence when fights involve larger gangs, but this is cleaner in the city-level regressions.

Step V: Other fighting parameters The estimation of the police attention cost δ, the average fixed cost from fighting $\bar{\gamma}$ and its dispersion σ_γ follows Rust (1987) Nested Fixed Point algorithm. Denote by $\Upsilon(\tau, \mathbf{V}) = \mathbf{F}^*$ a function that, for a given value of the parameters $\tau = [\delta, \bar{\gamma}, \sigma_\gamma]$ and matrix of violence $\mathbf{V}$, gives the vector of fight probabilities, which are the best responses of the all the dyads to the level of violence. The uniqueness of best response, discussed in section 4, plays a key role here. Denote by $K(\mathbf{F}) = \mathbf{V}$ the function that, for a given vector of fight probabilities, gives us the expected level of

[36] As shown there, fights–particularly those involving larger gangs–induce violence not only in the gangs' territories but also in adjacent areas. This is especially relevant in Naples, where districts are small and geographically contiguous. Using location-level variation would place spillover-affected districts in the control group, violating SUTVA and biasing estimates of the parameters downward. Aggregating to the city level mitigates this concern.

violence. Then, a fixed point of the NFXP is a pair $\{\tau^*, \mathbf{V}^*\}$ that satisfies:

i $\mathbf{V}^* = K^{-1}\Upsilon(\tau^*, \mathbf{V}^*)$

ii $\tau^* = \text{argmax}_\tau \sum_d \sum_t \Upsilon_{dt}(\tau, \mathbf{V}^*) I_{d_t}$

where I_{d_t} is an indicator for dyad d_t. The estimation routine begins by solving the fixed point mapping at an initial guess of the vector of parameters τ. To obtain the dyad d at time t best response function $\Upsilon_{dt}(\tau, \mathbf{V})$, I compute the profits under all possible scenarios (fight, peace, exit, etc.), which requires simulating police attention and demand. An equilibrium of the game is a fixed point of this algorithm. Once this has been found, the best response to this level of violence is computed, and these feeds into the log-likelihood which is maximized with respect to τ. This procedure is repeated until both expected violence and parameters converge. Informally, identification of the fixed cost comes from the average fighting probability and identification of the police cost comes from variation in fight propensity with the simulated police attention, given parameters estimated in step III and IV. The estimate of δ obtained with this procedure is sensible to the starting condition. For this reason, the estimation procedure is repeated on a grid of starting conditions, and the estimate with the minimum likelihood is chosen. Results are summarized in Figure A17.

Estimation results Table 2 shows the results of the estimation. The estimated price elasticity is negative and very small (a 1% increase in drug prices is associated with a 0.014% decline in market share), consistent with a very inelastic demand. This estimate should be interpreted cautiously given the coarse nature of the observed prices and their limited variation. Distance is also negatively related to demand, with a 1% increase in average hours associated with an estimated 0.412% decrease in quantity. Violence exerts a negative and sizable effect. At the average level of violence, a 1 (% at average violence) homicide is associated with a 0.75% (0.251%) decrease in demand–equivalent to a 18% increase in price, consistent with the evidence presented in Fact 3.

The estimated fighting ability of large gangs exceeds that of smaller gangs, corresponding to a 20% higher probability of winning a fight. The estimated threshold for fighting is 4.159, which is greater than the maximum difference in fighting abilities. This result is consistent with the high probability of draws, estimated at 77%. The estimated fighting shock is also large, at 1.74, indicating substantial unpredictability in fighting outcomes. For police reaction, at the average level of violence, one (% at average violence

Table 2: Estimated parameters

Variable	Notation	Estimate	S.e.	Elasticity
Step I: Demand parameters				
Price	α	-4.18e-06	9.29e-06	-0.014
Transportation cost	d	-0.000125	0.00005	-0.412
Violence disutility	μ	-0.011296	0.001160	-0.251
Step II: Fighting parameters				
Fighting ability large	ϕ_l	4.742		
Threshold fight	Φ	4.159		
Dispersion shock	σ_ϵ	1.738		
Step III: Police reaction parameters				
Police reaction	λ	0.006171	0.001530	0.151
Step IV: Violence parameters				
Violence (small)	θ_s	0.319044	0.155935	2.089
Violence (large)	θ_l	1.679729	0.717500	10.999
Spillover (small)	ψ_s	0.095152	0.086689	0.592
Spillover (large)	ψ_l	0.192055	0.087309	1.196
Step V: Other fighting parameters				
Police cost	δ	1.2218978		6.522
Average fight fixed cost	$\bar{\gamma}$	0.378146		38.658
Dispersion fight fixed cost	σ_γ	0.258544		

increase in) homicide is estimated to be associated with a 0.63% (0.15%) increase in police attention.

Consistent with the evidence documented in Fact 2, the estimated elasticities of violence to fights are large and heterogeneous. Fights involving only small gangs are associated with an approximate 2% increase in gang homicides, whereas those involving at least one large gang correspond to an 11% increase. Spillover effects to other locations are also sizable, with increases of 0.592% for fights involving only small gangs and 1.196% for those involving at least one large gang. Finally, police attention is estimated to impose a significant cost on criminal firms: at the mean level of enforcement, a 1% increase in police attention reduces average peace profits in the location by 6.5%. To get an idea of the magnitude, a fight involving a large gang is associated with a 10.75% decrease in average profits in the location. The average fixed cost of violence is also sizable, accounting for 38% of average total profits of a gang.

Model fit Figures A15 and A16 test the demand model fit by comparing, across areas, drugs, and years, the observed drug quantity share seized with that predicted by the model, both with and without Naples respectively. The fit is good, showing a strong positive correlation in both graphs. The correlation between standardized measures of quantity is approximately 67%. Table A18 reports the correlation between the observed share and the simulated share, controlling for drug, year, and city fixed effects. The correlation is positive and statistically different from zero in all specifications.

Table 3: Model fit - fights

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Data	Model	Random	Fixed	Data	Model	Random	Fixed	Data	Model	Random	Fixed
Average scale	-0.0916***	-0.149***	0.00396	-0.0150	-0.108***	-0.118	0.00332	0.0199	-0.0881**	-0.0680	0.00804	0.0109
	(0.0253)	(0.0444)	(0.0178)	(0.00961)	(0.0370)	(0.0834)	(0.0449)	(0.0215)	(0.0366)	(0.0801)	(0.0457)	(0.0230)
Number gangs	0.143***	0.112***	0.160***	0.414***	0.252***	0.0498	0.179***	0.460***	0.257***	0.110	0.169***	0.459***
	(0.0238)	(0.0351)	(0.0160)	(0.0176)	(0.0589)	(0.0789)	(0.0578)	(0.0274)	(0.0586)	(0.0721)	(0.0617)	(0.0280)
Observations	756	756	756	756	753	753	753	753	753	753	753	753
R2	0.139	0.146	0.111	0.763	0.365	0.602	0.213	0.806	0.382	0.657	0.217	0.807
P-value scale		0.247	0.001	0.006		0.896	0.082	0.001		0.795	0.138	0.012
P-value number		0.408	0.539	0.000		0.020	0.417	0.001		0.072	0.334	0.002
Location FE	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. Average scale is the average scale of gangs active in that location in that period. Columns (1), (5), (9) have as dependent variable a dummy indicating a fight in the data. Columns (2), (6), (10) have as dependent variable the sum of fighting probabilities predicted by the model. Columns (3), (7), (11) a dummy indicating a randomly predicted fight. Columns (4),(8),(12) the sum of randomly predicted fighting probabilities, randomly drawn from a normal distribution with the same moments as the predicted fighting probabilities. Dependent variables are standardized in all models. P-values indicate the equality of coefficients with the one for columns (1), (5), (9).

Figures A18 presents the average observed fight probability for five quintiles of model-predicted probability. The probability of fights rises across the first three quintiles and then declines slightly. Figures A19 and A20 compare correlations between observed and model-predicted fight probabilities and the number of gangs present (replicating Figure A31) and the average gang scale (replicating Figure A35). Model predictions closely track the empirical correlations, showing a positive correlation with the number of gangs–attenuated after the second quintile but positive overall, as confirmed in Table 3–and a negative correlation with average gang scale. In other words, the model is able to replicate the correlations shown in the descriptive evidence. Figures A21 and A22 extend this comparison to the standard deviation of gang scale (replicating Figure A40) and the maximum gang scale (replicating Figure A37), again showing close alignment between model predictions and data.

Table 3 provides a detailed analysis of these correlations, regressing the average gang scale and the number of gangs in a location on four dependent standardized variables: (1) an observed fight dummy (as in Table A2); (2) the sum of model-predicted fight probabilities in the location; (3) a dummy for a random fight generated using the average empirical fight probability; and (4) the sum of random fighting probabilities drawn from a normal distribution with the same moments as the predicted ones. The last two outcomes test whether the model is able to capture more than purely random fights. All specifications are estimated without fixed effects (columns 1–4), with location fixed effects (columns 5–8), and with location and year fixed effects (columns 9–12). The model does a good job replicating the main correlations in the data–clearly better than the random specifications. Table A19 replicates for the maximum, rather than average, scale of gangs in the location. Moreover, for other determinants of fights–such as the average distance of the location from others and the share of consumers–the model's predictions align closely with the data, as documented in Table A20.

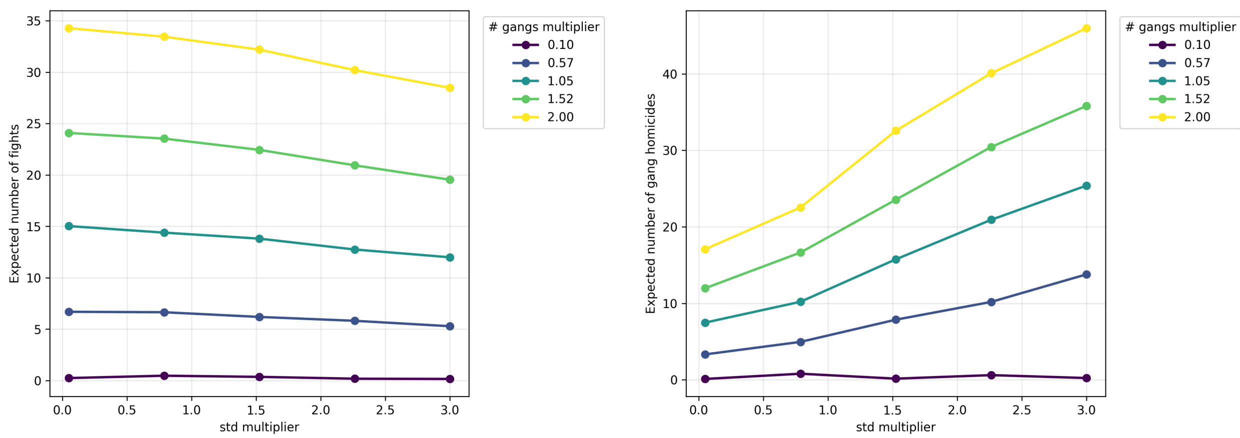
6. COUNTERFACTUALS

This section uses the estimated model to construct three sets of counterfactuals. First, it examines the impact of counterfactual changes in market structure on the expected number of fights and gang homicides, answering to the research question of the paper. Kingpin strategies are then evaluated. Second, it evaluates the effectiveness of selective enforcement. Third, it assesses two drug-legalization scenarios: (i) legalization of marijuana and (ii) joint legalization of marijuana and cocaine.

6.1. *Market structure, fragmentation, and kingpin strategies*

Market structure The first counterfactual examines how violence varies with market structure, addressing the main research question of the paper. To change the distribution of gangs, I randomly modify the ownership matrix–merging and splitting gangs–in the model while keeping the number of gangs constant. I summarize size inequality using the standard deviation of gangs' scale in the market. To vary the number of competitors, I randomly add or remove gangs while keeping the dispersion of gangs' scale constant. This procedure is repeated across different scenarios, using various multipliers of the observed number and standard deviation. For each scenario, the model's equilibrium is computed by solving for the fixed point described in Section 5. Since the model may admit multiple equilibria, the search for the fixed point is initialized from a belief of peace in the market, consistent with the estimation procedure.

Figure 2 presents the results separately for the expected number of fights (Panel A) and the expected number of homicides (Panel B). It is important to emphasize that this structural exercise is crucial for understanding the effects of both components of market structure on violence. In a reduced-form framework, it would be extremely difficult to obtain sufficient variation in one component while holding the other constant.



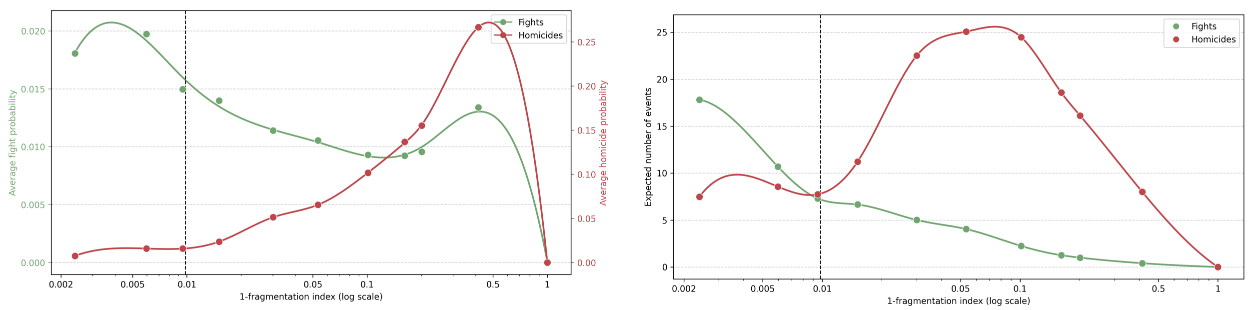
Panel A: Expected number of fights Panel B: Expected number of gang homicides

Notes: This figure presents the equilibrium expected number of fights (Panel A) and gang homicides (Panel B) for different levels of gang size standard deviation and different numbers of gangs. This counterfactual exercise is performed by merging or splitting gangs, while keeping the total number fixed, to change the distribution of gang sizes, and by adding or removing gangs, while keeping the distribution fixed, to change the total number of gangs.

Figure 2: Counterfactual - market structure

A counterfactual increase in inequality between gangs is associated with a decrease in the expected number of fights, and this effect is stronger when more gangs compete in the market. This result implies that the diseconomies of scale in the fighting mechanism–arising from spillovers and increased violence–outlined in Section 3 outweigh the economies of scale resulting in higher winning probabilities. To disentangle the contribution of each mechanism, Figure A23 replicates the simulations while sequentially muting different channels. Equalizing the probabilities of fighting outcomes further strengthens the negative relationship between inequality and the expected number of fights. As expected, equalizing the amount of violence generated in each fight, both within the affected locations and in neighboring ones, slightly weakens the relationship. Finally, shutting down spillovers entirely has the largest impact, indicating that they play the dominant role in driving this pattern. In contrast, a counterfactual increase in the number of competitors, holding their relative size distribution constant, increases both the expected number of fights and homicides.

Fragmentation To study the relationship between market fragmentation and the use of violence, I summarize the two components of market structure in an index of market fragmentation: $\sum_j (\frac{1}{Y} \sum_{y \in \mathcal{Y}_j} \frac{1}{N_y})^2$, where N_y is the number of gangs in location y. To reduce it, gangs are randomly split and new gangs are introduced. To increase it, gangs are randomly merged. I repeat this procedure for 11 scenarios ranging from very disaggregated markets to monopoly. Figure 3 presents the results.



Panel A: Average violence per dyad

Panel B: Violence in the whole market

Notes: This figure presents the average fight probability and the expected average number of homicides per dyad (Panel A) and for the overall market (Panel B), as predicted by the structural model across 11 simulated market structures, determining a fragmentation index ranging from approximately 0.002 to 1. The dashed line corresponds to the average fragmentation index in the province of Naples.

Figure 3: Counterfactual - market concentration

Panel A of Figure 3 plots how the average fighting probability (in green) and the average expected number of gang-homicides (in red) changes in each dyad. In a highly fragmented market the probability of fights is high. As concentration rises and a few large gangs emerge, the probability of fights increases even further. This is driven a reduction in the unpredictability of conflict. Further increases in market concentration are associated with a lower predicted fighting probability, reflecting rising costs from the diseconomies of scale in the fighting technology. At the other extreme of the distribution, the relationship reverses, as the expected benefits from fighting decreases non-linearly in the number of active gangs. In monopoly there is no violent competition, trivially. Multiplying the fighting probabilities by the expected number of gang homicides per fight yields an increasing curve in market concentration.

Panel B of Figure 3 shows how total violence in the province varies with market fragmentation. Given the previous results and the decline in fighting opportunities with the number of active gangs, the relationship between market concentration and the total number of fights is overall decreasing. For the expected number of gang homicides, the relationship is inverse U-shaped. This arises from the balance between a *declining probability* of fights and an *increasing intensity* of each fight. Interestingly, around the observed concentration, a plateau emerges, consistent with the null estimates reported in Table A2;

at this point, the declining probability of fights and the rising number of homicides per fight offset each other. Although the overall level of violence is similar at the two extremes of fragmentation, its nature differs: fragmented markets exhibit frequent, low-intensity fights, whereas concentrated markets feature infrequent, high-intensity conflicts.

The interpretation of results at very high levels of market concentration warrants caution, since the number of homicides generated by fights is estimated in relatively disaggregated markets. It is plausible that, at the upper end of the concentration distribution, homicides per fight could rise further, implying that total violence might continue to increase with market concentration. Therefore, it is possible that the overall relationship has two, rather the one, local maximums.

Anecdotally, this inverse U-shaped relationship between market structure and violence helps explaining worldwide differences. Cities with many small gangs, such as London or Naples, typically exhibit low levels of violence. By contrast, cities where several large gangs compete for territory, such as Chicago, Los Angeles or cities in Mexico, tend to have high gang homicides rates. At the other extreme, cities with highly concentrated markets–such as Palermo–are generally more peaceful.

Kingpin strategies These findings also help reconcile the seemingly contradictory evidence on the effectiveness of kingpin strategies (e.g. Moeller & Hesse, 2013; Vargas, 2014; Dell, 2015; Phillips, 2015; Burke, 2023). A kingpin strategy essentially fragments criminal groups. The simulations show that, in very fragmented markets, further fragmentation can reduce both crime and violence. In contrast, in oligopolistic markets at intermediate concentration levels, breaking up large criminal organizations can shift the market from a relatively peaceful equilibrium back toward a more violent one: a 1% decrease in concentration, in the middle of the distribution, is associated with a 5.5% increase in expected gang homicides.

6.2. Selective enforcement

Counterfactual Selective enforcement refers to sustained interventions targeting specific areas or groups (e.g., neighborhoods or drug hotspots), typically implemented through intensified policing and community engagement. In the simulation, this enforcement strategy is modeled by counterfactually altering the police's response to violent events, allowing us to capture different intensities of selective enforcement. Specifically, I simulate three degrees of reaction, thereby mimicking alternative scenarios of targeted enforcement intensity: (i) no reaction $\lambda = 0$, (ii) estimated $\hat{\lambda}$, and (iii) stronger reaction $\hat{\lambda} \times 2$.

Results Results are summarized in Figure A24. Panel A shows the average probability of fights per dyad across different levels of market concentration and police response. Stronger police reaction shifts the entire curve downward, indicating a lower average probability of fights. The effect is most pronounced in moderately concentrated markets, consistent with a higher elasticity of violence to fights. Panel B of the same figure confirms this pattern for market-level violence. Stronger police reaction not only lowers average fighting probabilities but also changes the shape of the relationship between market concentration and violence. Weaker police reaction is associated with higher violence, especially in highly concentrated markets, while stronger police reaction yields lower violence, particularly where market concentration is high. Doubling police reaction reduces expected homicides by about 14% at a fragmentation index near 1% and by about 59% near 10%.

These results suggest that, in moderately or highly concentrated markets, selective enforcement could be an effective strategy. A similar approach was implemented in Chicago, which was associated with a 23% reduction in shootings (Papachristos & Kirk, 2015).

6.3. Drug legalization

Counterfactual This final set of counterfactuals examines how drug legalization affects violent competition. I consider two policy experiments: (i) legalization of marijuana and (ii) joint legalization of marijuana and cocaine. Legalization is modeled as a zero price in the illicit sector, which collapses supply for the illegal product.[37] I replicate each scenario under three cost regimes: no fixed cost (the benchmark), a low fixed cost equal to the fifth percentile of gang peace profits, and a high fixed cost equal to the tenth percentile.

Results Figure A25 reports the results. Without fixed costs, legalization operates as a negative demand shock, lowering the returns to a larger market share. This *profitability* effect weakens incentives to fight and delivers an unambiguous decline in the probability of fights, the expected homicide rate, and market-wide violence (top row of Figure A25).

Fixed costs make the effects less clear-cut by introducing a *gambling-for-resurrection* effect: gangs may

[37] An alternative would be to introduce a parallel legal sector. I abstract from this for tractability, given the absence of price competition in the model and the lack of data on substitution between legal and illegal sources.

compete more aggressively to capture market share and cover the fixed cost. The net effect on violence reflects the balance between these opposing forces. The middle and bottom rows of Figure A25 present the results.

With fixed costs, the relationship between market concentration and violence remains broadly similar to the benchmark, but legalization alters the equilibrium shift. When only marijuana–a low-markup product–is legalized, violence still declines under both low and high fixed costs, but the magnitude of the decline is smaller and decreases as fixed costs rise. When both marijuana and cocaine are legalized, sharply reducing the profitability of market shares, the direction of the effect is no longer straightforward. At low levels of market concentration, the incremental market share from fighting is small, so the profitability effect dominates and legalization lowers violence. As concentration increases, however, the incremental market share from winning a fight grows. The gambling-for-resurrection effect strengthens relative to the profitability one and can overturn it. Therefore, in these markets legalization can unintentionally stimulate violent competition. The sharp fall in profitability induced by legalization, combined with the high incremental market share from winning, prompts gangs to gamble for survival, making them compete more aggressively and ultimately increases violence. In the simulations, legalization of both marijuana and cocaine, in the presence of "high" fixed costs, is associated with a 10.9% reduction in the expected number of gang homicides at 1% fragmentation and a 18% *increase* in gang homicides at 10% fragmentation.

Taken together, these counterfactual suggest that the effect of drug legalization on violence is not straightforward. In general, legalization reduces incentives for violent competition, but, in the presence of significant fixed costs, it may induce more aggressive conduct and, consequently, violence.

7. CONCLUSIONS

In this paper, I provide new evidence on the relationship between market structure, the number and size distribution of competitors, and violence in illicit markets, leveraging novel data on the retail drug market in Naples and a structural model. First, I present new facts on the fighting technology: (i) fights increase violence and police attention, heterogeneously by gang scale, (ii) fights reduce the number of active gangs, with stronger effects for larger gangs, and (iii) violence disrupts drug markets. Building on these facts, I characterize the trade-off gangs face when deciding whether to engage in violence.

To study how market structure shapes this trade-off, I develop a model of violent competition in which gangs decide whether to fight, consumers choose where to buy drugs, and police responds endogenously to violence. The model is estimated structurally and used to simulate counterfactual changes in market composition. The simulations show that greater size inequality reduces the expected number of fights, whereas an increase in the number of gangs raises it. In both cases, expected homicides rise. Overall, market fragmentation displays an inverse U-shaped relationship with expected homicides. This pattern reflects the balance between a declining probability of fights and an increasing intensity of each fight. Consequently, although the overall level of violence is similar at the two extremes of fragmentation, its nature differs: fragmented markets exhibit frequent, low-intensity fights, whereas concentrated markets feature infrequent, high-intensity conflicts.

I then evaluate three alternative enforcement strategies: (i) kingpin targeting, (ii) selective enforcement, and (iii) drug legalization. The effects of kingpin targeting depend critically on the initial market structure: in fragmented markets, removing dominant players may reduce violence, whereas in concentrated markets, fragmenting criminal groups can increase it. This mechanism helps reconcile seemingly contradictory empirical evidence on the effectiveness of kingpin policies. Selective enforcement increases the marginal cost of fighting and thereby reduces violence, particularly in concentrated markets where gangs are large, suggesting it may be an effective strategy in such contexts. Drug legalization operates as a negative demand shock and generally weakens incentives for violent competition. However, in the presence of fixed operating costs, it can trigger a gambling-for-resurrection effect, whereby struggling gangs undertake more aggressive violent competition to survive.

Overall, the main policy insight of the analysis is that the effectiveness of enforcement policies depends critically on underlying market structure: the same strategy may succeed, fail, or unintentionally escalate violence depending on the competitive environment.

This paper is an initial attempt to study violent competition in illicit markets and its consequences. It abstracts from coalition formation, intertemporal strategic behavior, and other forms of competition (such as price or quality), which future research could explore to provide a more comprehensive view of strategic interaction in illegal markets.

References

Aguirregabiria, V., & Ho, C.-Y. (2012). A dynamic oligopoly game of the us airline industry: Estimation and policy experiments. *Journal of Econometrics*, *168*(1), 156–173.

Aguirregabiria, V., & Vicentini, G. (2016). Dynamic spatial competition between multi-store retailers. *The Journal of Industrial Economics*, *64*(4), 710–754.

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

Angrist, J. D., & Kugler, A. D. (2008). Rural windfall or a new resource curse? coca, income, and civil conflict in colombia. *The Review of Economics and Statistics*, *90*(2), 191–215.

Baliga, S., & Sjöström, T. (2004). Arms races and negotiations. *The Review of Economic Studies*, *71*(2), 351–369.

Baliga, S., & Sjöström, T. (2012). The strategy of manipulating conflict. *American Economic Review*, *102*(6), 2897–2922.

Barbagallo, F. (2014). *Storia della camorra*. Gius. Laterza & Figli Spa.

Barton, J., Castillo, M., & Petrie, R. (2016). Negative campaigning, fundraising, and voter turnout: A field experiment. *Journal of Economic Behavior & Organization*, *121*, 99–113.

Bass, F. M., Krishnamoorthy, A., Prasad, A., & Sethi, S. P. (2005). Generic and brand advertising strategies in a dynamic duopoly. *Marketing Science*, *24*(4), 556–568.

Becker, G. S., Murphy, K. M., & Grossman, M. (2006). The market for illegal goods: the case of drugs. *Journal of Political Economy*, *114*(1), 38–60.

Becucci, S. (2004). Old and new actors in the italian drug trade: Ethnic succession or functional specialization? *European Journal on Criminal Policy and Research*, *10*, 257–283.

Benoit, J.-P. (1984). Financially constrained entry in a game with incomplete information. *The RAND Journal of Economics*, 490–499.

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica*, 889–917.

Bhatt, M. P., Heller, S. B., Kapustin, M., Bertrand, M., & Blattman, C. (2024). Predicting and preventing gun violence: An experimental evaluation of readi chicago. *The Quarterly Journal of Economics*, *139*(1), 1–56.

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

Bostanci, G., Yildirim, P., & Jerath, K. (2023). Negative advertising and competitive positioning. *Management Science*, *69*(4), 2361–2382.

Brancaccio, L., & Martone, V. (2019). The camorras in naples and campania: Business, groups and families. In *Italian mafias today* (pp. 30–46). Edward Elgar Publishing.

Brown, Z. Y., Montero, E., Schmidt-Padilla, C., & Sviatschi, M. M. (2025). Market structure and extortion: Evidence from 50,000 extortion payments. *Review of Economic Studies*, *92*(3), 1595–1624.

Bruhn, J., et al. (2021). Competition in the black market: Estimating the causal effect of gangs in chicago.

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

Burke, P. J. (2023). Reducing organized criminal violence through leadership removals: evidence from the drug wars on chicago's westside, 2010-2019. *Journal of crime and justice*, *46*(1), 102–123.

Campaniello, N., Gray, R., & Mastrobuoni, G. (2016). Returns to education in criminal organizations: Did going to college help michael corleone? *Economics of Education Review*, *54*, 242–258.

Castillo, J. C., Mejía, D., & Restrepo, P. (2020). Scarcity without leviathan: The violent effects of cocaine supply shortages in the mexican drug war. *Review of Economics and Statistics*, *102*(2), 269–286.

Catino, M. (2014). How do mafias organize?: Conflict and violence in three mafia organizations. *European Journal of Sociology*, *55*(2), 177–220.

Catino, M. (2019). *Mafia organizations*. Cambridge University Press.

Ciliberto, F., & Tamer, E. (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, *77*(6), 1791–1828.

Clarke, D., & Tapia Schythe, K. (2023). Eventdd: Stata module to panel event study models and generate event study plots.

Coate, S. (2004). Political competition with campaign contributions and informative advertising. *Journal of the European Economic Association*, *2*(5), 772–804.

Cook, P. J., Ludwig, J., Venkatesh, S., & Braga, A. A. (2007). Underground gun markets. *The Economic Journal*, *117*(524), F588–F618.

Corsaro, N., Hunt, E. D., Hipple, N. K., & McGarrell, E. F. (2012). The impact of drug market pulling levers policing on neighborhood violence: An evaluation of the high point drug market intervention. *Criminology & Public Policy*, *11*(2), 167–199.

Cruz, I. L., & Torrens, G. (2023). Hidden drivers of violence diffusion: Evidence from illegal oil siphoning in mexico. *Journal of Economic Behavior & Organization*, *206*, 26–70.

DCSA. (2024). *Relazione annuale direzione centrale servizi antridroga 2024* (Tech. Rep.).

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

DIA. (2015). *Relazione del ministreo dell'interno al parlamento: attività svolta e risultati conseguiti dalla direzione investigativa antimafia* (Tech. Rep.).

DIA. (2020). *Relazione del ministreo dell'interno al parlamento: attività svolta e risultati conseguiti dalla direzione investigativa antimafia* (Tech. Rep.).

DIA. (2022). *Relazione del ministreo dell'interno al parlamento: attività svolta e risultati conseguiti dalla direzione investigativa antimafia* (Tech. Rep.).

Dickenson, M. (2014). The impact of leadership removal on mexican drug trafficking organizations. *Journal of Quantitative Criminology*, *30*(4), 651–676.

Dixit, A. (1980). The role of investment in entry-deterrence. *The Economic Journal*, *90*(357), 95–106.

Donohue, J. J., & Levitt, S. D. (1998). Guns, violence, and the efficiency of illegal markets. *The American Economic Review*, *88*(2), 463–467.

Dugato, M., Calderoni, F., & Berlusconi, G. (2020). Forecasting organized crime homicides: Risk terrain modeling of camorra violence in naples, italy. *Journal of interpersonal violence*, *35*(19-20), 4013–4039.

EMCDDA. (2013). *European monitoring centre for drugs and drug addiction: Drug report* (Tech. Rep.).

Erfle, S., & McMillan, H. (1990). Media, political pressure, and the firm: The case of petroleum pricing in the late 1970s. *The Quarterly Journal of Economics*, *105*(1), 115–134.

Eurispes. (2005). *Rapporto italia 2005* (Tech. Rep.).

Fox, A. M., Katz, C. M., Choate, D. E., & Hedberg, E. C. (2015). Evaluation of the phoenix truce project: a replication of chicago ceasefire. *Justice Quarterly*, *32*(1), 85–115.

Fudenberg, D., & Tirole, J. (1986). A theory of exit in duopoly. *Econometrica*, 943–960.

Galenianos, M., & Gavazza, A. (2017). A structural model of the retail market for illicit drugs. *American Economic Review*, *107*(3), 858–896.

Galenianos, M., Pacula, R. L., & Persico, N. (2012). A search-theoretic model of the retail market for illicit drugs. *The Review of Economic Studies*, *79*(3), 1239–1269.

Gambetta, D., & Reuter, P. (2017). Conspiracy among the many: the mafia in legitimate industries. In

Transnational organized crime (pp. 247–266). Routledge.

Goldberg, P. K. (1995). Product differentiation and oligopoly in international markets: The case of the us automobile industry. *Econometrica*, 891–951.

Gribaudi, M. G., et al. (2009). Clan camorristi a napoli: radicamento locale e traffici internazionali. In *Traffici criminali. camorra, mafie e reti internazionali dell'illegalità* (pp. 187–240). Bollati Boringhieri.

Grossman, G. M., & Shapiro, C. (1984). Informative advertising with differentiated products. *The Review of Economic Studies*, *51*(1), 63–81.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.

Klement, C., & Blokland, A. (2023). Preventing outlaw biker crime in the netherlands or just changing the dark figure? estimating the impact of the dutch whole-of-government approach on outlaw biker crime using interrupted time series analysis. *European Journal of Criminology*, *20*(4), 1187–1210.

Le Moglie, M., & Sorrenti, G. (2022). Revealing "mafia inc."? financial crisis, organized crime, and the birth of new enterprises. *Review of Economics and Statistics*, *104*(1), 142–156.

Levitt, S. D. (2017). The economics of crime. *Journal of Political Economy*, *125*(6), 1920–1925.

Levitt, S. D., & Venkatesh, S. A. (2000). An economic analysis of a drug-selling gang's finances. *The Quarterly Journal of Economics*, *115*(3), 755–789.

Lind, J. T., Moene, K. O., & Willumsen, F. (2014). Opium for the masses? conflict-induced narcotics production in afghanistan. *Review of Economics and Statistics*, *96*(5), 949–966.

Lindo, J. M., & Padilla-Romo, M. (2018). Kingpin approaches to fighting crime and community violence: evidence from mexico's drug war. *Journal of health economics*, *58*, 253–268.

Marrazzo, G. (2005). *Il camorrista*. Tullio Pironti Editore.

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

Mazzeo, M. J. (2002). Product choice and oligopoly market structure. *RAND Journal of Economics*, 221–242.

Melnikov, N., Schmidt-Padilla, C., & Sviatschi, M. M. (2022). Gangs, labor mobility and development. *Econometrica*.

Milgrom, P., & Roberts, J. (1982). Limit pricing and entry under incomplete information: An equilibrium analysis. *Econometrica*, 443–459.

Mirenda, L., Mocetti, S., & Rizzica, L. (2022). The economic effects of mafia: firm level evidence. *American Economic Review*, *112*(8), 2748–2773.

Moeller, K., & Hesse, M. (2013). Drug market disruption and systemic violence: Cannabis markets in copenhagen. *European Journal of Criminology*, *10*(2), 206–221.

Monnier, M. (1862). *La camorra notizie storiche raccolte e documentate*. G. Barbera.

Monteiro, J., & Rocha, R. (2017). Drug battles and school achievement: evidence from rio de janeiro's favelas. *Review of Economics and Statistics*, *99*(2), 213–228.

Papachristos, A. V., & Kirk, D. S. (2015). Changing the street dynamic: Evaluating chicago's group violence reduction strategy. *Criminology & Public Policy*, *14*(3), 525–558.

Phillips, B. J. (2015). How does leadership decapitation affect violence? the case of drug trafficking organizations in mexico. *The Journal of Politics*, *77*(2), 324–336.

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

Polo, M. (1995). Internal cohesion and competition among criminal organizations. *The economics of organised crime*, 87–115.

Reiss, P., & Bresnahan, T. (1990). Entry in monopoly markets. *Review of Economic Studies*, *57*(4), 531–53.

Reuter, P. (1985). *The organization of illegal markets: An economic analysis*.

Reuter, P. (2014). Drug markets and organized crime. *The Oxford handbook of organized crime*, 359–381.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 999–1033.

Sales, I. (1994). *La camorra, le camorre*. Editori Riuniti.

Sales, I. (2022). *Storia delle camorre: passato e presente*. Rubbettino.

Schelling, T. C. (1967). Economics and criminal enterprise. *The Public Interest*, *7*, 61.

Schmalensee, R. (1978). Entry deterrence in the ready-to-eat breakfast cereal industry. *The Bell Journal of Economics*, 305–327.

Schultz, C. (2007). Strategic campaigns and redistributive politics. *The Economic Journal*, *117*(522), 936–963.

Spence, A. M. (1977). Entry, capacity, investment and oligopolistic pricing. *The Bell Journal of Economics*, 534–544.

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

Transcrime, C. (2013). *Progetto pon sicurezza 2007–2013. gli investimenti delle mafie* (Tech. Rep.).

Vargas, R. (2014). Criminal group embeddedness and the adverse effects of arresting a gang's leader: a comparative case study. *Criminology*, *52*(2), 143–168.

APPENDIX

A. PROOFS AND DERIVATIONS

A.1. Derivation demand

Consumer problem Consumer solves:

$$\max_{d_{(y,j)}} \sum_{(y,j)} d_{(y,j)} \mathbb{E} u_{ijyvt} \quad \forall v, t$$
$$\text{s.t.} \sum_{(y,j)} d_{(y,j)} = 1$$

Since in this model an individual is defined by product specific shocks $(\epsilon_{ijvt}, \epsilon_{iyvt})$, this implictly defines the set of individual attributes that lead to the purchase of good (j, y). Formally:

$$A_{jy} = \{(\epsilon_{ijvt}, \epsilon_{iyvt}) | \mathbb{E} u_{ijyvt} \geq \mathbb{E} u_{i\tilde{j}\tilde{y}vt} \forall \tilde{j} \in \mathcal{J} \text{ and } \forall \tilde{y} \in \mathcal{Y}\}$$

Assuming ties occur with zero probability, and the iid nature of the two EVT1 shocks, the market share of the (j, y) bundle is just an integral over the mass of consumers in region A_{jyv}:

$$s_{jyvt} = \int_{A_{jy}} dP(\epsilon_{ijvt}, \epsilon_{iyvt}) =^{iid} \int_{A_{jy}} dP(\epsilon_{ijvt}) P(\epsilon_{iyvt})$$

I will now work to find an explicit expression definying the set A_{jy}. In particular, I will exploit the nested structure of the model. Notice that U can decompose:

$$P(ivt \in A_{jy}) = P_{ivt \to jy} = P_{ivt \to j|y} \times P_{ivt \to y}$$

In the remaining of this section I will derive an expression for these two probabilities.

Within nests Let's start with $P_{ivt \to j|y}$. A consumer in area y chooses the organization which gives him the highest utility. As standard in the literature, denote by $\delta_{jyvt} = \alpha_v p_{vt} + \mu_v V_{yt} + \xi_y + \xi_{vt} + \xi_{yvt}$ the mean consumer valuation. Formally:

$$P_{ivt \to j|y} = P(u_{ijyvt} \geq u_{i\tilde{j}yvt} \ \forall \tilde{j} \in \mathcal{J}_y) = P(\epsilon_{i\tilde{j}vt} \leq \epsilon_{ijvt} + \delta_{jyvt} - \delta_{\tilde{j}yvt} \ \forall \tilde{j} \in \mathcal{J}_y)$$

Since the ϵ_s are independent, this set cumulative distribution can be rewritten as the product of the individual cumulative distributions:

$$P_{ivt \to j|y} | \epsilon_{ijvt} = \prod_{\tilde{j} \in y - j} \exp[-\exp[\epsilon_{ijvt} + \delta_{jyvt} - \delta_{\tilde{j}yvt}]]$$

The problem with the expression is that I don't know ϵ_{ijvt}. So I integrate it away:

$$P_{ivt\to j|y} = \int_{-\infty}^{+\infty}[\prod_{\tilde{j}\in y-j} e^{-e^{-[\epsilon_{ijvt}+\delta_{jyvt}-\delta_{\tilde{j}yvt}]}}]e^{-\epsilon_{ijvt}}e^{e^{-\epsilon_{ijvt}}}d\epsilon_{ijvt} = \frac{\exp(\delta_{jyvt})}{\sum_{\tilde{j}\in y}\exp(\delta_{\tilde{j}yvt})}$$

To conclude, notice that the mean consumer valuation is identical for all gangs in the same location. Therefore:

$$P_{ivt\to j|y} = \frac{1}{N_{j\in y}}$$

Across nests Now I need to compute the second probability: $P_{ivt\to y}$. The maximum utility from nest/area y is:

$$\begin{cases} \max_{j\in y} u_{ijyvt} & y \neq 0 \\ 1 & y = 0 \end{cases}$$

where the second case comes from the normalization of expected utility of outside option. Substituting our equation for the indirect utility:

$$\max_{j\in y}\delta_{jyvt}+\epsilon_{iyvt}+\epsilon_{ijvt}+d_v g_{iy} = (\max_{j\in y}\epsilon_{ijvt})+\mu_v V_{yt}+\alpha_v p_{vt}+\xi_y+\xi_{vt}+\xi_{yvt}+\epsilon_{iyvt}+d_v g_{iy}$$

Due to the extreme value properties of the ETV1 component I have (approximation of the expectation of the maximum of an ETV1 distribution):

$$\mathbb{E}(\max_{j\in y}\epsilon_{ijvt}) \approx \gamma$$

Substituting into the previous problem:

$$(\max_{j\in y} u_{ijyvt}) \approx \alpha_v p_{vt} + \mu_v V_{yt} + \xi_y + \xi_{vt} + \xi_{yvt} + \epsilon_{iyvt} + d_v g_{iy} + \gamma$$

Hence the maximum expected indirect utility is:

$$V_{iyvt} = \bar{\delta}_{yvt} + \epsilon_{iyvt} + d_v g_{iy} + \gamma$$

Similarly to the reasoning done before, the consumer chooses the nest which gives him the highest expected utility:

$$P_{ivt\to y} = P(V_{iyvt} \geq V_{i\tilde{y}vt}\ \forall\tilde{y} = 0,1,...,Y) = \frac{\exp(\bar{\delta}_{yvt}+d_v g_{iy})}{1+\sum_{\tilde{y}\in\mathcal{Y}}\exp(\bar{\delta}_{\tilde{y}vt}+d_v g_{i\tilde{y}})}$$

Before turning to the unconditional demand, we will now see how to transform this across nest share into an unconditional demand for drugs in area y. The unconditional demand for drugs in nest y can be written as weighted sum of the conditional demands:

$$s_{yvt} = \sum_{y_i} \frac{M_{y_i}}{\sum_y M_y} s_{yvt|y_i} = \sum_{y_i} \frac{M_{y_i}}{\sum_y M_y} \frac{\exp(\bar{\delta}_{yvt} + d_v g_{iy})}{1 + \sum_{\tilde{y} \in \mathcal{Y}} \exp(\bar{\delta}_{\tilde{y}vt} + d_v g_{i\tilde{y}})}$$

Denote by $S_{vt} = 1 + \sum_{\tilde{y} \in \mathcal{Y}} \exp(\bar{\delta}_{\tilde{y}vt} + d_v g_{i\tilde{y}})$. Taking logs:

$$log(s_{yvt}) = log(\sum_{y_i} \frac{M_{y_i}}{\sum_y M_y} s_{yvt|y_i}) = -log(S_{vt}) + \bar{\delta}_{yvt} + log(\sum_{y_i} \frac{M_{y_i}}{\sum_y M_y} exp(d_v g_{iy}))$$

Which, by adding the fixed effects, becomes equation 2:

$$log(s_{yvt}) = \mu_v V_{yt} + f_y + f_{vt} + \xi_{yvt}$$

Unconditional demand Finally, we can put together the two probabilities to compute the probability of consumer i choosing the bundle (j, y) for drug v at t:

$$P(ivt \in A_{jy}) = P_{ivt \to jy} = \frac{1}{N_y} \times \frac{\exp(\bar{\delta}_{yvt} + d_v g_{iy})}{1 + \sum_{\tilde{y} \in \mathcal{Y}} \exp(\bar{\delta}_{\tilde{y}vt} + d_v g_{i\tilde{y}})}$$

Demand is represented by this set of individual probabilities:

$$P_{ivt \to jy} = \frac{\exp(\bar{\delta}_{yvt} + d_v g_{iy})}{N_y \sum_{\tilde{j} \in y} [1 + \sum_{\tilde{y} \in \mathcal{Y}} \exp(\bar{\delta}_{\tilde{y}vt} + d_v g_{i\tilde{y}})]} \tag{9}$$

To conclude, we need to compute the market shares. Returning to our first derivations we have:

$$s_{jyvt} = \int_{A_{jy}} dP(\epsilon_{ijvt}) P(\epsilon_{iyvt}) \approx \sum_{y_i} \frac{M_{y_i}}{\sum_y M_y} \frac{\exp(\bar{\delta}_{yvt} + d_v g_{iy})}{N_y \sum_{\tilde{j} \in y} [1 + \sum_{\tilde{y} \in \mathcal{Y}} \exp(\bar{\delta}_{\tilde{y}vt} + d_v g_{i\tilde{y}})]}$$

These individual buying probabilities can be used to compute the aggregate demanded quantities.

A.2. Equilibria and conditional choice probabilities

Fighting game with no uncertainty:

$$\begin{cases} \phi_j > \Phi + \phi_o & p(F, NF) + (1-p)(F, NF) \to \text{o exits} \to F_{j,o} = 0 \\ -\Phi < \phi_j - \phi_o \leq \Phi \ \& \ \pi_j^{NF} > \pi_j^d \ \& \ \pi_o^{NF} > \pi_o^d & p(F, NF) + (1-p)(NF, F) \to \textbf{ME} \to F_{j,o} = 0 \\ -\Phi < \phi_j - \phi_o \leq \Phi \ \& \ \pi_j^{NF} > \pi_j^d \ \& \ \pi_o^{NF} \leq \pi_o^d & p(NF, F) + (1-p)(NF, F) \to \text{j exits} \to F_{j,o} = 0 \\ -\Phi < \phi_j - \phi_o \leq \Phi \ \& \ \pi_j^{NF} \leq \pi_j^d \ \& \ \pi_o^{NF} > \pi_o^d & p(F, NF) + (1-p)(F, NF) \to \text{o exits} \to F_{j,o} = 0 \\ -\Phi < \phi_j - \phi_o \leq \Phi \ \& \ \pi_j^{NF} \leq \pi_j^d \ \& \ \pi_o^{NF} \leq \pi_o^d & p(NF, F) + (1-p)(F, NF) \to \textbf{ME} \to F_{j,o} = 0 \\ \phi_o > \Phi + \phi_j & p(NF, F) + (1-p)(NF, F) \to \text{j exits} \to F_{j,o} = 0 \end{cases}$$

Fighting game with uncertainty:

$$\begin{cases} E\pi_j - \pi_j^p \leq E\pi_j - \pi_j^{NF} \leq \gamma_j \;\&\; E\pi_o - \pi_o^p \leq E\pi_o - \pi_o^{NF} \leq \gamma_o & p(F,NF) + (1-p)(NF,F) \rightarrow \textbf{ME} \rightarrow \textbf{ME} \rightarrow F_{j,o} = 0 \\ E\pi_j - \pi_j^p \leq \gamma_j \leq E\pi_j - \pi_j^{NF} \;\&\; E\pi_o - \pi_o^p \leq E\pi_o - \pi_o^{NF} \leq \gamma_o & p(F,NF) + (1-p)(F,NF) \rightarrow \textbf{ME} \rightarrow \text{o exits} \rightarrow F_{j,o} = 0 \\ \gamma_j \leq E\pi_j - \pi_j^p \leq E\pi_j - \pi_j^{NF} \;\&\; E\pi_o - \pi_o^p \leq E\pi_o - \pi_o^{NF} \leq \gamma_o & p(F,NF) + (1-p)(F,NF) \rightarrow \textbf{ME} \rightarrow \text{o exits} \rightarrow F_{j,o} = 0 \\ E\pi_j - \pi_j^p \leq E\pi_j - \pi_j^{NF} \leq \gamma_j \;\&\; E\pi_o - \pi_o^p \leq \gamma_o \leq E\pi_o - \pi_o^{NF} & p(NF,F) + (1-p)(NF,F) \rightarrow \textbf{ME} \rightarrow \text{j exits} \rightarrow F_{j,o} = 0 \\ E\pi_j - \pi_j^p \leq \gamma_j \leq E\pi_j - \pi_j^{NF} \;\&\; E\pi_o - \pi_o^p \leq \gamma_o \leq E\pi_o - \pi_o^{NF} & p(NF,F) + (1-p)(F,NF) \rightarrow \text{nothing happens} \rightarrow F_{j,o} = 0 \\ \gamma_j \leq E\pi_j - \pi_j^p \leq E\pi_j - \pi_j^{NF} \;\&\; E\pi_o - \pi_o^p \leq \gamma_o \leq E\pi_o - \pi_o^{NF} & p(F,F) + (1-p)(F,NF) \rightarrow \textbf{ME} \rightarrow \text{fight} \rightarrow F_{j,o} = 1 \\ E\pi_j - \pi_j^p \leq E\pi_j - \pi_j^{NF} \leq \gamma_j \;\&\; \gamma_o \leq E\pi_o - \pi_o^p \leq E\pi_o - \pi_o^{NF} & p(NF,F) + (1-p)(NF,F) \rightarrow \textbf{ME} \rightarrow \text{j exits} \rightarrow F_{j,o} = 0 \\ E\pi_j - \pi_j^p \leq \gamma_j \leq E\pi_j - \pi_j^{NF} \;\&\; \gamma_o \leq E\pi_o - \pi_o^p \leq E\pi_o - \pi_o^{NF} & p(NF,F) + (1-p)(F,F) \rightarrow \textbf{ME} \rightarrow \text{fight} \rightarrow F_{j,o} = 1 \\ \gamma_j \leq E\pi_j - \pi_j^p \leq E\pi_j - \pi_j^{NF} \;\&\; \gamma_o \leq E\pi_o - \pi_o^p \leq E\pi_o - \pi_o^{NF} & p(F,F) + (1-p)(F,F) \rightarrow \text{fight} \rightarrow \text{fight} \rightarrow F_{j,o} = 1 \end{cases}$$

Conditional choice probabilities:

$$\begin{cases} Pr(NF,NF) = Pr(\pi_j^{NF} \leq E\pi_j \leq \pi_j^p \;\&\; \pi_o^{NF} \leq E\pi_o \leq \pi_o^p), \\ Pr(F,F) = Pr(\pi_j^{NF} \leq \pi_j^p \leq E\pi_j \;\&\; \pi_o^{NF} \leq E\pi_o \leq \pi_o^p; \\ \qquad \pi_j^{NF} \leq E\pi_j \leq \pi_j^p \;\&\; \pi_o^{NF} \leq \pi_o^p \leq E\pi_o; \pi_j^{NF} \leq \pi_j^p \leq E\pi_j \;\&\; \pi_o^{NF} \leq \pi_o^p \leq E\pi_o), \\ Pr(NF,F) = Pr(E\pi_j \leq \pi_j^{NF} \leq \pi_j^p \;\&\; \pi_o^{NF} \leq E\pi_o \leq \pi_o^p; E\pi_j \leq \pi_j^{NF} \leq \pi_j^p \;\&\; \pi_o^{NF} \leq \pi_o^p \leq E\pi_o) + \\ \qquad \int T(NF,F) 1[(E\pi_j, E\pi_o) \in R] dF_{E\pi_j, E\pi_o}, \\ Pr(F,NF) = Pr(\pi_j^{NF} \leq E\pi_j \leq \pi_j^p \;\&\; E\pi_o \leq \pi_o^{NF} \leq \pi_o^p; \pi_j^{NF} \leq \pi_j^p \leq E\pi_j \;\&\; E\pi_o \leq \pi_o^{NF} \leq \pi_o^p) + \\ \qquad \int (1 - T(NF,F)) 1[(E\pi_j, E\pi_o) \in R] dF_{E\pi_j, E\pi_o} \end{cases}$$

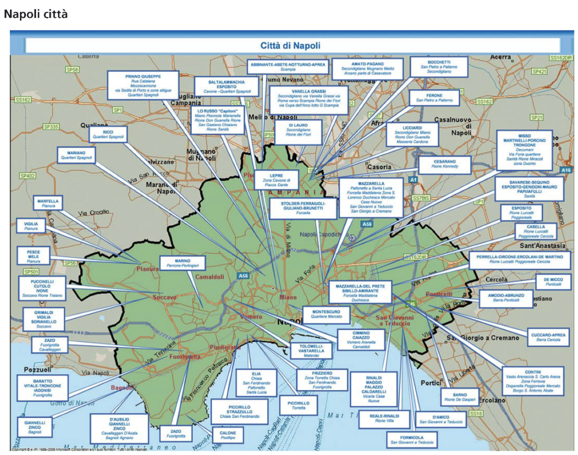
$$R = \{E\pi_j, E\pi_o | E\pi_j \leq \pi_j^{NF} \leq \pi_j^p \;\&\; E\pi_o \leq \pi_o^{NF} \leq \pi_o^p\}$$

The first two CCPs are straightforward. For example, the model predicts *(NF,NF)* uniquely if and only if the expected profits fall in the central quadrant. Equalities three and four are more involved. Both include the cases where *(NF,F)* or *(F,NF)* is the unique equilibrium of the game (top-left and bottom-right quadrants), and also the cases where these outcomes are potentially observable in the multiple-equilibria region and are selected. Following Ciliberto & Tamer, 2009, define $T(NF,F)$ as the selection mechanism. This function partitions the multiple-equilibria quadrant into the two possible equilibria, thus determining the choice probabilities.

Outcome probabilities:

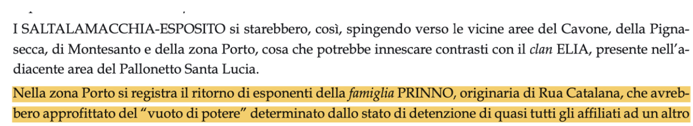
$$
\begin{cases}
Pr(F_{jo}=1) = Pr(\pi_j^{NF} \leq \pi_j^p \leq E\pi_j \;\&\; \pi_o^{NF} \leq E\pi_o \leq \pi_o^p; \\
\qquad \pi_j^{NF} \leq E\pi_j \leq \pi_j^p \;\&\; \pi_o^{NF} \leq \pi_o^p \leq E\pi_o; \pi_j^{NF} \leq \pi_j^p \leq E\pi_j \;\&\; \pi_o^{NF} \leq \pi_o^p \leq E\pi_o), \\
Pr(F_{jo}=0) = Pr(\pi_j^{NF} \leq E\pi_j \leq \pi_j^p \;\&\; \pi_o^{NF} \leq E\pi_o \leq \pi_o^p; \\
\qquad E\pi_j \leq \pi_j^{NF} \leq \pi_j^p \;\&\; \pi_o^{NF} \leq E\pi_o \leq \pi_o^p; E\pi_j \leq \pi_j^{NF} \leq \pi_j^p \;\&\; \pi_o^{NF} \leq \pi_o^p \leq E\pi_o; \\
\qquad \pi_j^{NF} \leq E\pi_j \leq \pi_j^p \;\&\; E\pi_o \leq \pi_o^{NF} \leq \pi_o^p; \pi_j^{NF} \leq \pi_j^p \leq E\pi_j \;\&\; E\pi_o \leq \pi_o^{NF} \leq \pi_o^p; \\
\qquad E\pi_j \leq \pi_j^{NF} \leq \pi_j^p \;\&\; E\pi_o \leq \pi_o^{NF} \leq \pi_o^p)
\end{cases}
$$

B. ADDITIONAL FIGURES



Notes: The figure presents an example of map reported in the DIA report.

Figure A1: DIA report - map ← Return to text

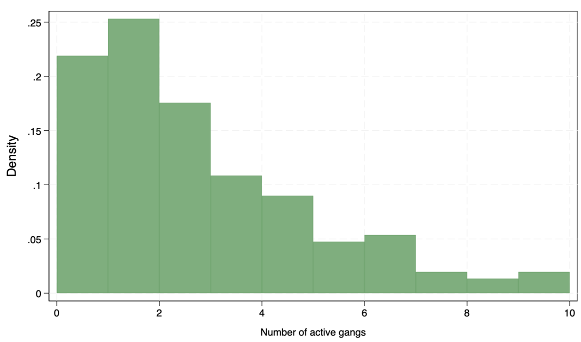


I SALTALAMACCHIA-ESPOSITO si starebbero, così, spingendo verso le vicine aree del Cavone, della Pignasecca, di Montesanto e della zona Porto, cosa che potrebbe innescare contrasti con il *clan* ELIA, presente nell'adiacente area del Pallonetto Santa Lucia.

Nella zona Porto si registra il ritorno di esponenti della *famiglia* PRINNO, originaria di Rua Catalana, che avrebbero approfittato del "vuoto di potere" determinato dallo stato di detenzione di quasi tutti gli affiliati ad un altro

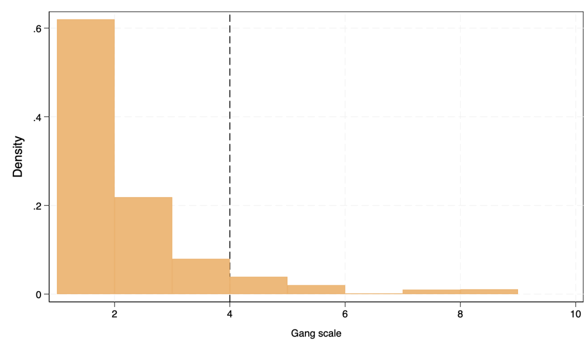
Notes: The figure presents an example of text reported in the DIA report. The highlighted text describes an entry of one gang in one location.

Figure A2: DIA report - text ← Return to text



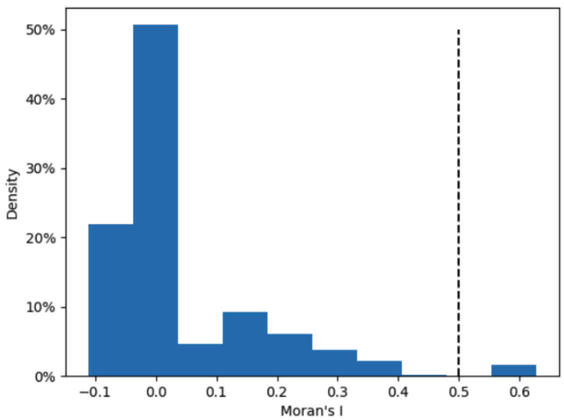
Notes: The figure shows the distribution of maximum gang scale, where gang scale is defined as the number of locations in which a gang operates in a given year. Each observation represents the scale of each gang-year in the sample period.

Figure A3: Density number active gangs ← Return to text



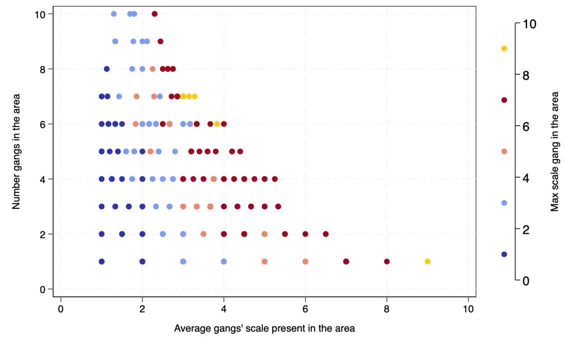
Notes: The figure shows the distribution of maximum gang scale, where gang scale is defined as the number of locations in which a gang operates in a given year. Each observation represents the scale of each gang-year in the sample period.

Figure A4: Density gang scale ← Return to text



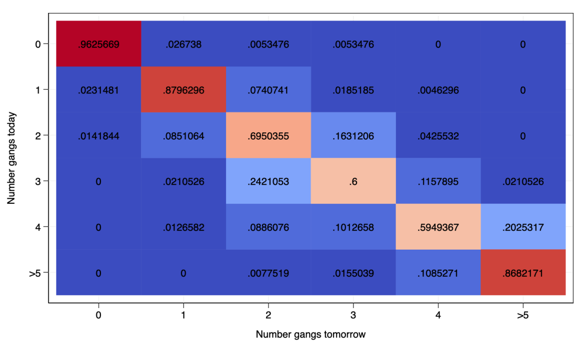
Notes: This figure presents the Moran's I index of spatial concentration of gangs spatial distribution. The vertical line indicates 0.5, the rule of thumb for spatial concentration.

Figure A5: Moran's I ← Return to text



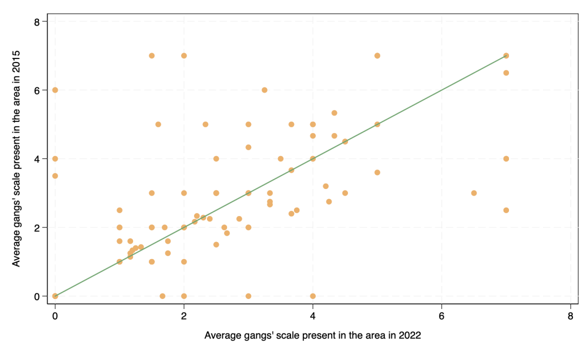
Notes: The figure shows cross-sectional variation in gang presence across location-year observations. A location is a district within Naples, and a city in the province.

Figure A6: Cross section variation ← Return to text



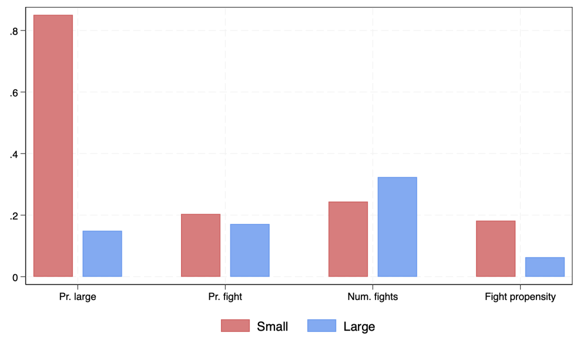
Notes: The figure shows the transition matrix of the number of gangs active in a given location. A location is a district within Naples, and a city in the province.

Figure A7: Transition matrix number of gangs ← Return to text



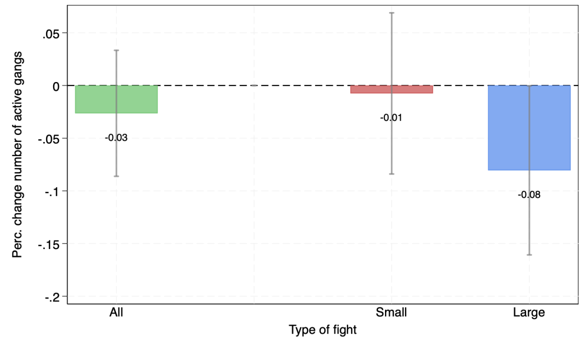
Notes: The figure shows average gang scale in 2022 plotted against average gang scale in 2015 for each location. The green line indicates the 45-degree line where gang scale remains unchanged between the two years. A location is a district within Naples, and a city in the province.

Figure A8: Time variation in gang scale ← Return to text



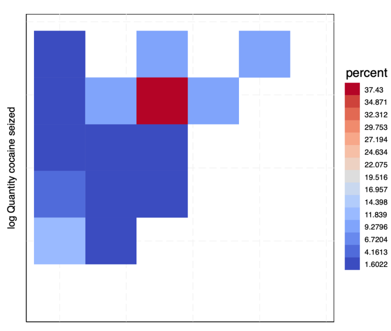
Notes: The figure shows the average of different metrics of violence for small and large gangs. In the first two columns, the proportion of fights involving small/large gangs. In the second two, the probability of a gang being involved in a fight in a given year. Then, the number of fights a gang is involved in a given year. Finally, the fighting propensity for the two categories of gangs. This is simply the number of fights over the number of areas where the gang is present in that year. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province.

Figure A9: Number, probability and propensity of fights per gang size ← Return to text



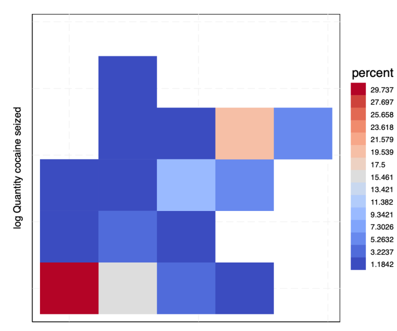
Notes: The figure shows the percentage change in the number of active gangs in the location between the (last continuous) fighting period and the first following peaceful one. 90% confidence interval shown. Three set of fights are presented: all fights, fights involving only small gangs, fights involving at least one large gang. For these last two categories, the other type of fight is eliminated from the sample. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province.

Figure A10: Change in number of gangs following fights ← Return to text



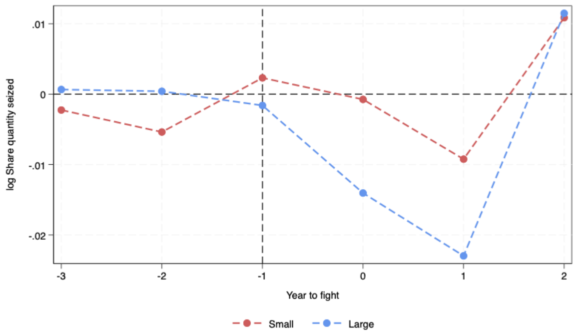
Notes: The figure shows the percent of location-years for log quantity of cocaine seized and log quantity of heroin seized. Locations are districts within Naples and cities in the province.

Figure A11: Heatplot - seizures cocaine and heroin ← Return to text



Notes: The figure shows the percent of location-years for log quantity of cocaine seized and log quantity of heroin seized. Locations are districts within Naples and cities in the province.

Figure A12: Heatplot - seizures cocaine and marijuana ← Return to text



Notes: The figure presents the log share of quantity seized (+1)–defined as the ratio of the quantity of a drug seized in a city to the total quantity of that drug seized in the province during the same year–for observations relative to the first fight in each city, separately for small and large fights. Each observation corresponds to a city–drug–year. The sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (the 90th percentile of the scale distribution). The scale of a gang in a given year is the number of locations in which it is present. A fight is defined as large if at least one large gang is involved.

Figure A13: Seized quantity by years to first fight

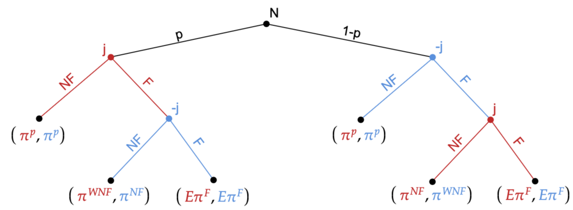
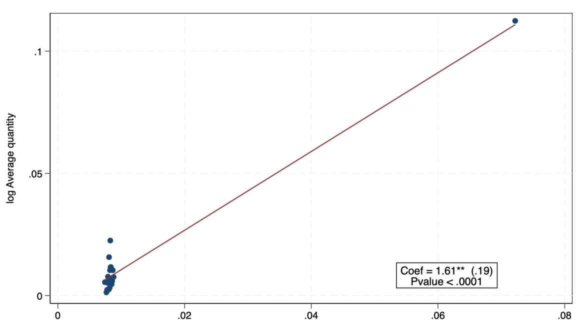
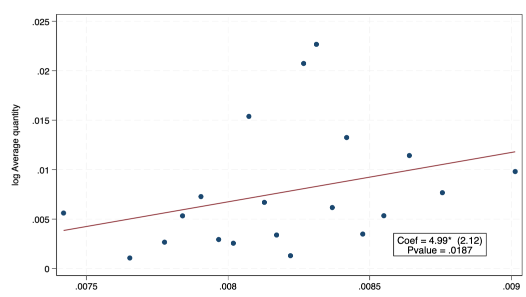


Figure A14: Fighting game ← Return to text



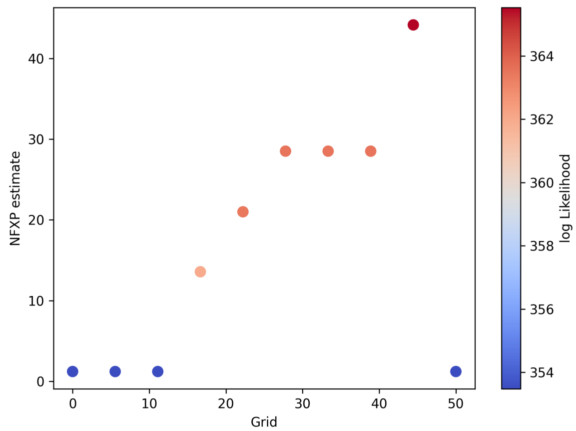
Notes: This figure presents scatter plot and linear fit of the log average share of drug seized in an area, in a given year, for a given drug, and the log average share predicted by the model.

Figure A15: Model fit - quantities ← Return to text



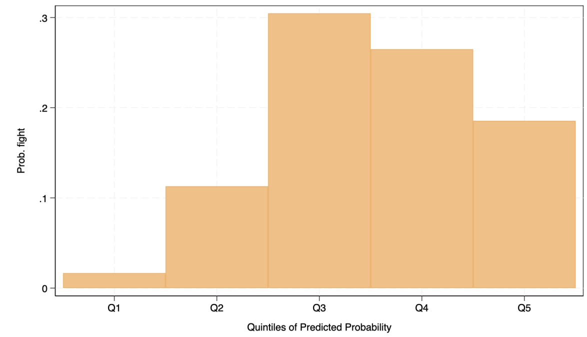
Notes: This figure presents scatter plot and linear fit of the log average share of drug seized in an area, in a given year, for a given drug, and the log average share predicted by the model. We exclude the city of Naples from the sample.

Figure A16: Model fit - quantities, no Naples ← Return to text



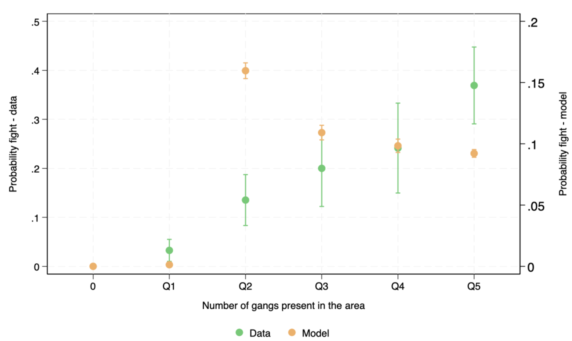
Notes: The figure presents the likelihood of each NFXP estimate starting from different points in the δ grid. Please refer to section 5.

Figure A17: NFXP likelihood grid ← Return to text



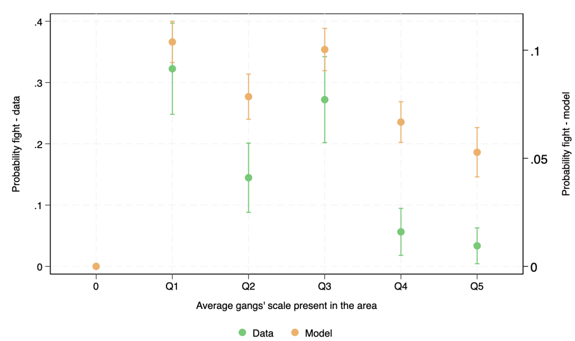
Notes: This figure presents the average probability of a fight occurring in a location–year for each of the five quintiles of the sum of predicted fight probabilities in that observation.

Figure A18: Correlation prob. fight and predicted probability ← Return to text



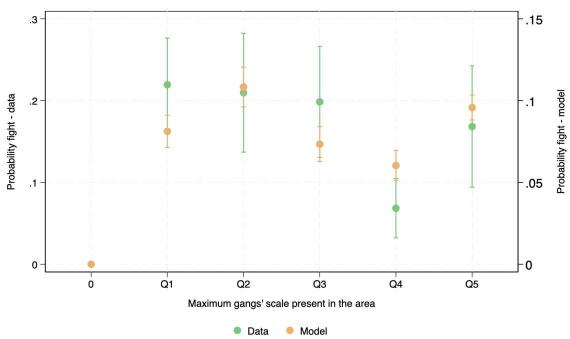
Notes: This figure presents the average probability of gang homicides for locations-years, in the data and as predicted by the model, with no gangs, and for the five quintiles of the number of gangs present in the area distribution. An homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. The scale of a gang, in a given year, is the number of locations in which it is present.

Figure A19: Correlation gang homicide and number of gangs ← Return to text



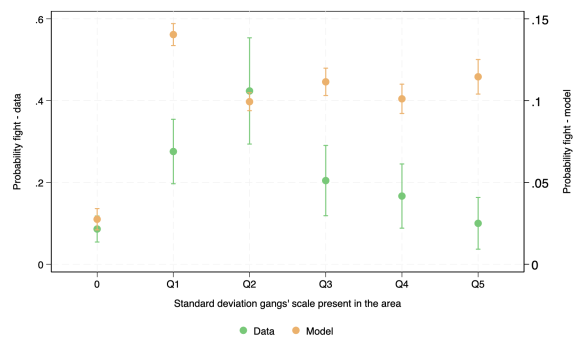
Notes: This figure presents the average probability of gang homicides for locations-years, in the data and as predicted by the model, with no gangs, and for the five quintiles of the average of scale of gangs present in the area distribution. An homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. The scale of a gang, in a given year, is the number of locations in which it is present.

Figure A20: Correlation gang homicide and average gang scale ← Return to text



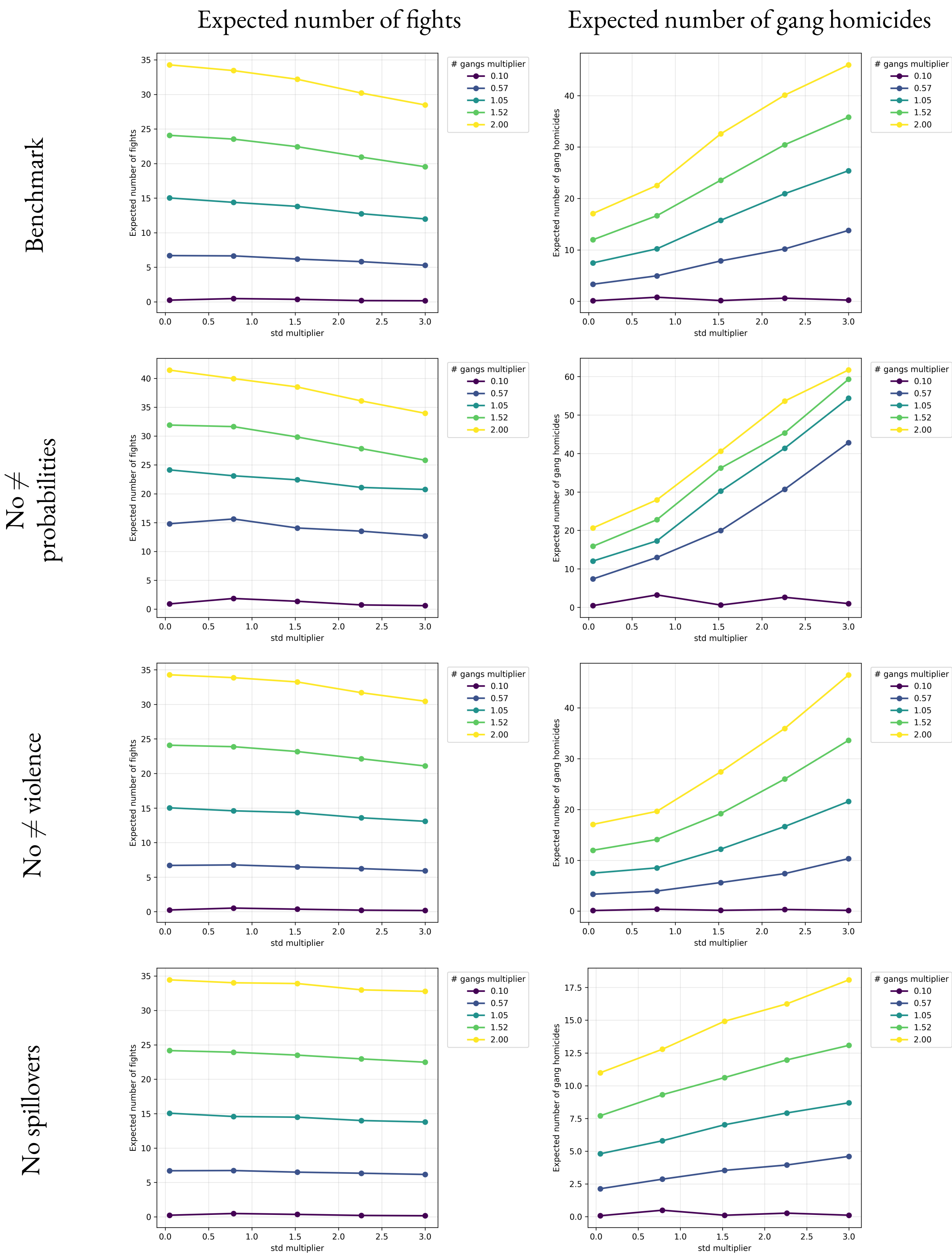
Notes: This figure presents the average probability of gang homicides for locations-years, in the data and as predicted by the model, with no gangs, and for the five quintiles of the maximum scale of gangs present in the area distribution. An homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. The scale of a gang, in a given year, is the number of locations in which it is present.

Figure A21: Correlation gang homicide and max gang scale ← Return to text



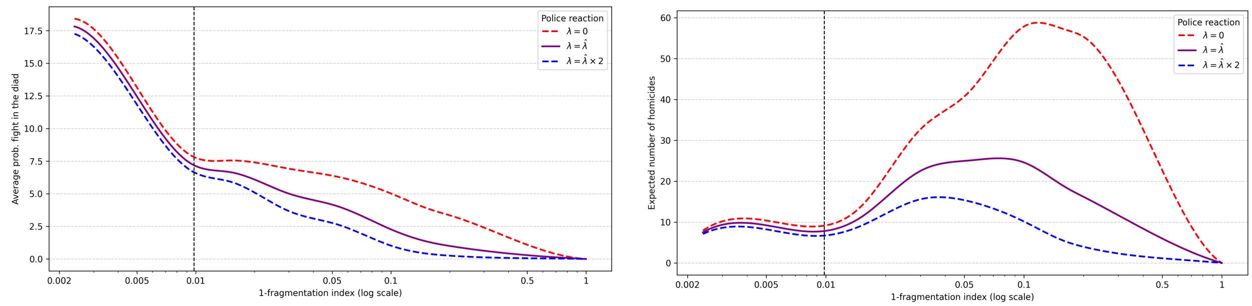
Notes: This figure presents the average probability of gang homicides for locations-years, in the data and as predicted by the model, with no gangs, and for the five quintiles of the standard deviation of scale of gangs present in the area distribution. An homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. The scale of a gang, in a given year, is the number of locations in which it is present.

Figure A22: Correlation gang homicide and sd gang scale ← Return to text



Notes: This figure presents the equilibrium expected number of fights (Panel A) and gang homicides (Panel B) for different levels of gang size standard deviation and different numbers of gangs. This counterfactual exercise is performed by merging or splitting gangs, while keeping the total number fixed, to change the distribution of gang sizes, and by adding or removing gangs, while keeping the distribution fixed, to change the total number of gangs. In the first row, fighting outcome probabilities are equalized between gangs of different scale. In the second row, the number of gang homicides and related spillovers are equalized between gangs of different scale. In the third row, spillovers are muted.

Figure A23: Counterfactual – marker structure: mechanisms ← Return to text



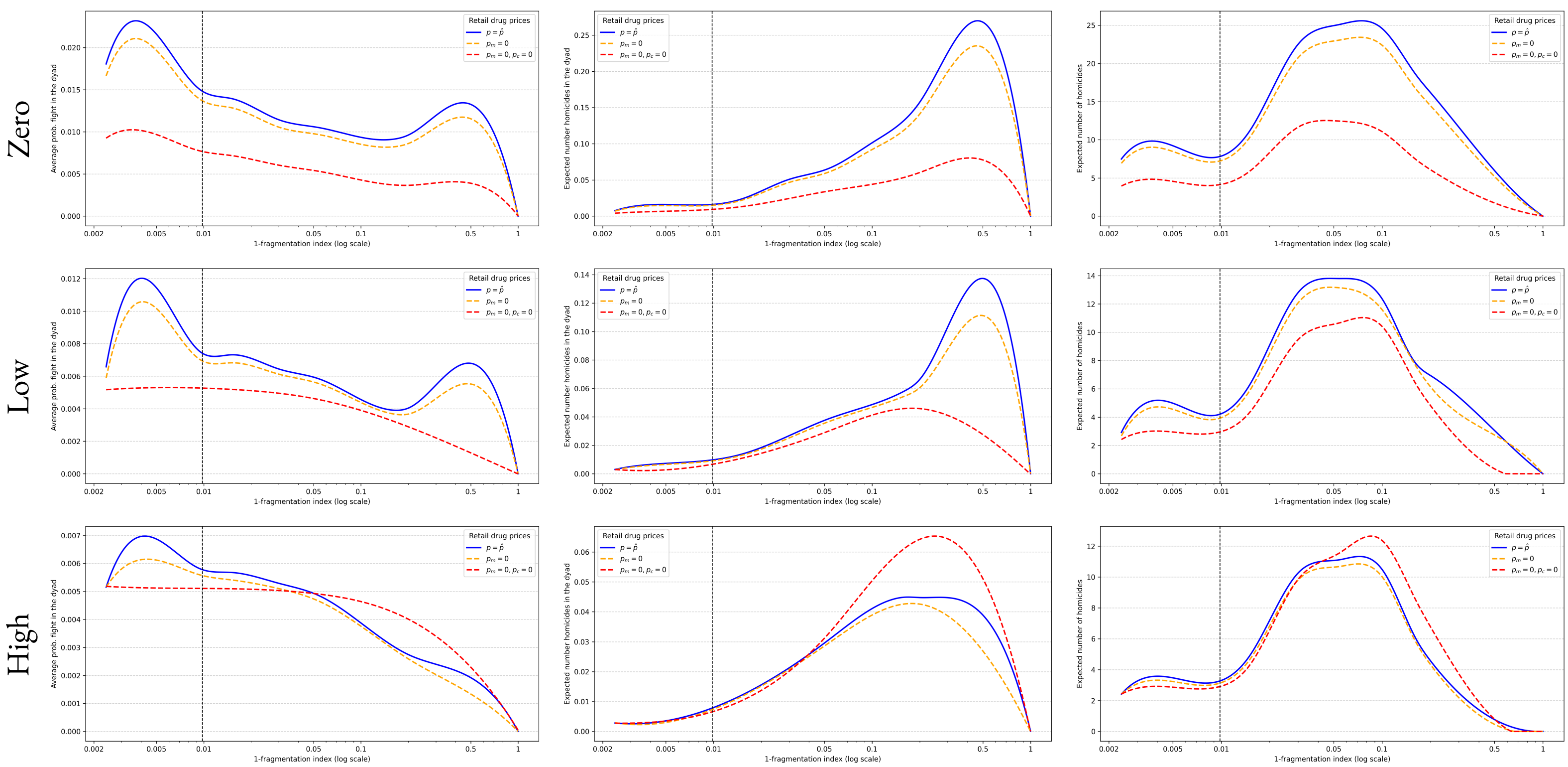
Panel A: Average violence per dyad

Panel B: Violence in the whole market

Notes: This figure presents the average fight probability per dyad (Panel A) and the expected number of homicides in the market (Panel B), as predicted by the structural model for different levels of police reaction $\lambda = 0$, $\lambda = \hat{\lambda}$, $\lambda = 2 \times \hat{\lambda}$. The predictions are obtained by simulating the model across 11 market structure scenarios–determining an fragmentation index ranging from approximately 0.002 to 1–and then fitting a curve through these simulated outcomes. Please refer to section 6.

Figure A24: Counterfactual - selective enforcement ← Return to text

Pr. fight per dyad | Expected violence per dyad | Violence in the market



Notes: This figure presents the average fight probability per dyad (first column), the expected number of homicides per dyad (second column), and the expected number of homicides in the market (third column), as predicted by the structural model in three counterfactual scenarios: (i) prices equal to the observed ones, (ii) price of marijuana set to 0, (ii) prices of both marijuana and cocaine set to 0. The first row reports simulations with no fixed costs (benchmark model); the second row, with low fixed costs (equal to the fifth percentile of gang peace profits); and the third row, with high fixed costs (equal to the tenth percentile of gang peace profits). The predictions are obtained by simulating the model across 11 market structure scenarios–determining a fragmentation index ranging from approximately 0.002 to 1–and then fitting a curve through these simulated outcomes.

Figure A25: Counterfactual – drug legalization ← Return to text

C. ADDITIONAL TABLES

Table A1: Fights and location characteristics ← Return to text

Dep. Variable: Fight	(1)	(2)	(3)	(4)	(5)	(6)
Number gangs	0.0619***					
	(0.00899)					
Average scale		-0.0579***				
		(0.0110)				
Standard deviation scale			-0.0675***			
			(0.0167)			
Maximum scale				-0.0188**		
				(0.00906)		
Share consumers					0.0492**	
					(0.0213)	
Average distance						-0.0323*
						(0.0177)
Observations	756	756	511	756	756	756
Mean Dep	0.0245	0.2264	0.2982	0.2264	0.1000	0.1004

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is a dummy indicating a fight in the observation. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. Average scale is the average scale of gangs in the location, measured as the number of locations in which the gang is present. Standard deviation scale is the standard deviation of scale of gangs present in the observation. Maximum scale their maximum. Share of consumers is the share of estimated consumers in the location. Average distance is the average distance in km of the location from all the others. Mean dependent is reported in the following way: column (1) is the average fight probability in locations with 1 gang, column (2) in locations with average scale equal to 1, column (3) standard deviation equal to 0, column (4) maximum scale equal to 1, column (5) share consumers below the first quartile of the distribution, column (6) average distance below the first quartile of the distribution.

Table A2: Average scale and violence ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: probability fight						
Average scale	-0.0579***		-0.0341***	-0.0549***		-0.0328**
	(0.0110)		(0.00942)	(0.0165)		(0.0136)
Number gangs		0.0619***	0.0533***		0.102***	0.0955***
		(0.00899)	(0.00883)		(0.0222)	(0.0218)
Observations	756	756	756	753	753	753
R-squared	0.0517	0.0444	0.0720	0.318	0.326	0.333
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.226	0.226	0.226	0.226	0.226	0.226
Mean Dep | Number gangs = 1	0.024	0.024	0.024	0.024	0.024	0.024
Panel B: probability fight, max scale						
Maximum scale	-0.0188**		-0.0286***	-0.00777		-0.0251**
	(0.00906)		(0.00721)	(0.0135)		(0.0120)
Number gangs		0.0619***	0.0667***		0.102***	0.108***
		(0.00899)	(0.00948)		(0.0222)	(0.0220)
Observations	756	756	756	753	753	753
R-squared	0.0167	0.0444	0.0715	0.306	0.326	0.330
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Maximum scale = 1	0.226	0.226	0.226	0.226	0.226	0.226
Panel C: probability homicide						
Average scale	-0.0296***		-0.0145	-0.0344**		-0.0320*
	(0.0104)		(0.00971)	(0.0172)		(0.0172)
Number gangs		0.0374***	0.0337***		0.0169	0.0105
		(0.00815)	(0.00816)		(0.0224)	(0.0224)
Observations	756	756	756	753	753	753
R-squared	0.011	0.033	0.035	0.251	0.249	0.251
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.301	0.301	0.301	0.301	0.301	0.301
Panel D: probability gang homicide						
Average scale	-0.0341***		-0.0190**	-0.0149		-0.0143
	(0.00942)		(0.00883)	(0.0252)		(0.0250)
Number gangs		0.0386***	0.0338***		0.00522	0.00237
		(0.00816)	(0.00846)		(0.0213)	(0.0207)
Observations	756	756	756	753	753	753
R-squared	0.022	0.051	0.057	0.267	0.267	0.267
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.245	0.245	0.245	0.245	0.245	0.245

	(1)	(2)	(3)	(4)	(5)	(6)
Panel D: number gang homicide						
Average scale	-0.0722***		-0.0438**	-0.0546		-0.0469
	(0.0227)		(0.0198)	(0.0481)		(0.0488)
Number gangs		0.0746***	0.0636***		0.0426	0.0332
		(0.0177)	(0.0167)		(0.0530)	(0.0536)
Observations	756	756	756	753	753	753
R-squared	0.0270	0.0513	0.0601	0.354	0.353	0.355
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.283	0.283	0.283	0.283	0.283	0.283
Panel E: number gang homicide per 1000 individuals						
Average scale	-0.000995		-0.000358	-0.00245		-0.00211
	(0.000814)		(0.000732)	(0.00187)		(0.00181)
Number gangs		0.00151**	0.00142***		0.00188	0.00146
		(0.000590)	(0.000532)		(0.00201)	(0.00194)
Observations	756	756	756	753	753	753
R-squared	0.0027	0.011	0.011	0.231	0.230	0.232
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.008	0.008	0.008	0.008	0.008	0.008
Panel F: probability non-gang homicide						
Average scale	0.00451		0.00449	-0.0195		-0.0176
	(0.00692)		(0.00747)	(0.0199)		(0.0199)
Number gangs		-0.00117	-0.0000349		0.0116	0.00813
		(0.00465)	(0.00504)		(0.0130)	(0.0128)
Observations	756	756	756	753	753	753
R-squared	0.000558	0.0000664	0.000558	0.159	0.158	0.160
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.057	0.057	0.057	0.057	0.057	0.057
Panel G: number non-gang homicide						
Average scale	-0.0205*		-0.0155	-0.0286		-0.0341
	(0.0114)		(0.0131)	(0.0263)		(0.0293)
Number gangs		0.0152	0.0113		-0.0168	-0.0236
		(0.00921)	(0.0105)		(0.0341)	(0.0366)
Observations	756	756	756	753	753	753
R-squared	0.00436	0.00425	0.00645	0.212	0.211	0.213
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.170	0.170	0.170	0.170	0.170	0.170

Notes: HDFE linear regressions. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. Average scale is the average scale of gangs in the location, measured as the number of locations in which the gang is present. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide.

Table A3: Sensitivity definition scale ← Return to text

	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dep. Variable:	Fight	Fight	Fight	Fight	Dep. Variable:	Fight	Fight	Fight	Fight
Panel A: number of locations (benchmark measure)					**Panel C: sum of locations weighted by consumers share**				
Average scale	-0.0861***		-0.0514**		Average scale	-0.0867***		-0.0518**	
	(0.0259)		(0.0213)			(0.0260)		(0.0214)	
Maximum scale				-0.0562**	Maximum scale				-0.0577**
				(0.0269)					(0.0269)
Number gangs		0.102***	0.0955***	0.108***	Number gangs		0.102***	0.0956***	0.108***
		(0.0222)	(0.0218)	(0.0220)			(0.0222)	(0.0218)	(0.0220)
Panel B: sum of population					**Panel D: sum of locations weighted by quantity seized share**				
Average scale	-0.101***		-0.0665**		Average scale	-0.0898***		-0.0593**	
	(0.0314)		(0.0265)			(0.0317)		(0.0230)	
Maximum scale				-0.0439	Maximum scale				-0.0543*
				(0.0324)					(0.0299)
Number gangs		0.102***	0.0943***	0.107***	Number gangs		0.102***	0.0960***	0.106***
		(0.0222)	(0.0215)	(0.0224)			(0.0222)	(0.0215)	(0.0217)
Observations	753	753	753	753	Observations	753	753	753	753
Location FE	Yes	Yes	Yes	Yes	Location FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Year FE	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. In Panel A average scale is the average scale of gangs in the location, measured as the number of locations in which the gang is present. In Panel B scale is measured as the (standardized) total number of people living in the locations where the gang is present. In Panel C scale is measured as the number of locations where the gang is present, weighted by the average share of estimated consumers across drugs. Independent variables are standardized to easy comparison.

Table A4: Effects of fights, unit-year ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Panel A: number gang homicides																				
Fight	0.522***			0.527***							0.290***			0.334***						
	(0.119)			(0.116)							(0.0867)			(0.0830)						
Fight small		0.456***			0.476***		0.470***		0.477***			0.209**			0.272***		0.213**		0.256***	
		(0.125)			(0.124)		(0.126)		(0.126)			(0.0827)			(0.0803)		(0.0868)		(0.0852)	
Fight large			0.575**			0.567**		0.654**		0.645**			0.362*			0.365**		0.489**		0.488**
			(0.256)			(0.238)		(0.265)		(0.253)			(0.187)			(0.164)		(0.212)		(0.198)
Observations	753	753	753	753	753	753	724	656	724	656	753	753	753	753	753	753	706	608	706	608
R-squared	0.377	0.356	0.341	0.404	0.385	0.368	0.387	0.283	0.409	0.306	0.342	0.332	0.333	0.375	0.364	0.360	0.366	0.286	0.389	0.305
Mean Dep | Fight = 0	0.109	0.141	0.207	0.109	0.141	0.207	0.109	0.109	0.109	0.109	0.107	0.144	0.205	0.107	0.144	0.205	0.107	0.107	0.107	0.107
Panel B: number arrests																				
Fight	0.847			0.629							0.966			1.254						
	(0.823)			(0.789)							(1.080)			(1.068)						
Fight small		0.0580			0.278		0.125		0.288			0.480			0.985		0.443		1.102	
		(0.697)			(0.689)		(0.717)		(0.698)			(0.974)			(0.913)		(1.038)		(0.963)	
Fight large			2.946			1.515		3.288		1.614			2.308			1.419		2.365		1.391
			(2.085)			(2.022)		(2.185)		(2.091)			(2.255)			(2.151)		(2.805)		(2.688)
Observations	753	753	753	753	753	753	724	656	724	656	753	753	753	753	753	753	706	608	706	608
R-squared	0.827	0.827	0.828	0.849	0.848	0.849	0.832	0.827	0.854	0.846	0.827	0.827	0.828	0.849	0.849	0.849	0.833	0.836	0.854	0.857
Mean Dep | Fight = 0	13.945	14.417	14.985	13.945	14.417	14.985	13.945	13.945	13.945	13.945	13.301	13.918	14.817	13.301	13.918	14.817	13.301	13.301	13.301	13.301
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	No	No	Yes	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes
Control	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed	Peace	Peace	Peace	Peace	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed	Peace	Peace	Peace	Peace
Including year after	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Sample is restricted to observations with at least one active gang. Fight is a dummy variable equal to one if a fight is recorded for that observation. Fight small is a dummy variable equal to one if a fight is recorded and the gang is small. Fight large is a dummy variable equal to one if a fight is recorded and the gang is large. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. In Panel B, the outcome variable is the sum of arrests for drug and conspiracy causes. In columns 7,9,17,19 we exclude from the sample observations with fights involving large gangs. In columns 8,10,18,20 we exclude from the sample observations with fights involving small gangs. In columns 11 to 20 the fight dummies are equal to one also the period after the fight outbreak. Arrests are divided by 30–number of districts–for Naples city, since we uniquely observe arrests at the city level.

Table A5: Fights and spillovers within Naples ← Return to text

	Fight	No fight	Difference
Panel A: All			
Large - Within Naples	0.480	0.396	+0.083*
Large - Outside Naples	0.109	0.096	+0.013
Small - Within Naples	0.406	0.421	-0.015
Small - Outside Naples	0.103	0.101	+0.002
Panel B: Conditional on peace			
Large - Within Naples	0.165	0.139	+0.026
Large - Outside Naples	0.074	0.060	+0.014
Small - Within Naples	0.147	0.149	-0.001
Small - Outside Naples	0.066	0.064	+0.002
Panel C: Gangs present in Naples			
Large	0.434	0.381	+0.054
Small	0.412	0.396	+0.021
Panel D: Gangs present in Naples, conditional on peace			
Large	0.145	0.122	+0.022
Small	0.151	0.142	0.009
Panel E: Gangs ever present in Naples			
Large	0.485	0.377	+0.101*
Small	0.423	0.415	+0.008
Panel F: Gangs ever present in Naples, conditional on peace			
Large	0.168	0.121	+0.048
Small	0.154	0.149	+0.005

Notes: Each observation is a gang-location-drug-year. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Table presents the mean number of gang homicides in locations within and outside Naples for gang-years in which the gang is involved in at least one fight, but the gang was not present in that location, by gang size. Locations are districts within Naples and cities in the province. An homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide.

Table A6: Effects of fights, city-year ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Panel A: number gang homicides																				
Fight	0.476***			0.521***							0.339***			0.415***						
	(0.148)			(0.147)							(0.101)			(0.101)						
Fight small		0.127			0.194		0.368***		0.377***			0.0407			0.155		0.258***		0.290***	
		(0.269)			(0.231)		(0.132)		(0.131)			(0.229)			(0.181)		(0.0844)		(0.0872)	
Fight large			1.818**			1.725**		1.043*		0.932*			0.717**			0.582*		0.717*		0.578*
			(0.842)			(0.760)		(0.592)		(0.551)			(0.351)			(0.314)		(0.385)		(0.345)
Observations	525	525	525	525	525	525	512	483	512	483	525	525	525	525	525	525	504	458	504	458
R-squared	0.702	0.698	0.711	0.711	0.706	0.717	0.676	0.712	0.685	0.720	0.701	0.697	0.700	0.710	0.706	0.707	0.254	0.698	0.290	0.706
Mean Dep | Fight = 0	0.095	0.281	0.162	0.095	0.281	0.162	0.095	0.095	0.095	0.095	0.088	0.265	0.149	0.088	0.265	0.149	0.088	0.088	0.088	0.088
Panel B: number arrests																				
Fight	0.383			1.360							1.132			2.388						
	(1.456)			(1.670)							(1.640)			(1.877)						
Fight small		-6.471			-5.463		-0.421		0.239			-7.317			-6.043		0.660		1.365	
		(6.030)			(5.760)		(1.132)		(1.219)			(7.970)			(7.393)		(1.479)		(1.393)	
Fight large			30.48			29.63		4.636		4.363			14.49			13.82		3.330		2.932
			(23.28)			(22.52)		(6.759)		(6.786)			(12.48)			(12.05)		(6.151)		(5.894)
Observations	525	525	525	525	525	525	512	483	512	483	525	525	525	525	525	525	504	458	504	458
R-squared	0.980	0.981	0.982	0.981	0.981	0.982	0.979	0.983	0.980	0.984	0.980	0.981	0.981	0.981	0.982	0.982	0.786	0.992	0.801	0.992
Mean Dep | Fight = 0	9.734	20.068	12.818	9.734	20.068	12.818	9.734	9.734	9.734	9.734	9.377	18.897	11.434	9.377	18.897	11.434	9.377	9.377	9.377	9.377
Panel C: number gang homicides, no Naples																				
Fight	0.476***			0.493***							0.339***			0.379***						
	(0.148)			(0.151)							(0.101)			(0.104)						
Fight small		0.368***			0.389***		0.368***		0.374***			0.258***			0.316***		0.258***		0.290***	
		(0.132)			(0.130)		(0.132)		(0.130)			(0.0844)			(0.0866)		(0.0844)		(0.0872)	
Fight large			1.043*			1.050*		1.043*		1.030*			0.717*			0.685*		0.717*		0.681*
			(0.592)			(0.579)		(0.592)		(0.574)			(0.385)			(0.374)		(0.385)		(0.365)
Observations	517	517	517	517	517	517	510	477	510	477	517	517	517	517	517	517	504	453	504	453
R-squared	0.278	0.242	0.261	0.307	0.271	0.288	0.259	0.258	0.285	0.286	0.256	0.230	0.245	0.292	0.266	0.268	0.254	0.250	0.290	0.272
Mean Dep | Fight = 0	0.096	0.111	0.127	0.096	0.111	0.127	0.096	0.096	0.096	0.096	0.088	0.108	0.125	0.088	0.108	0.125	0.088	0.088	0.088	0.088
Panel D: number arrests, no Naples																				
Fight	0.383			0.952							1.132			1.630						
	(1.457)			(1.549)							(1.640)			(1.658)						
Fight small		-0.421			0.274		-0.421		0.253			0.660			1.348		0.660		1.365	
		(1.132)			(1.223)		(1.132)		(1.213)			(1.479)			(1.405)		(1.479)		(1.393)	
Fight large			4.636			4.460		4.636		4.762			3.330			3.005		3.330		3.026
			(6.760)			(6.646)		(6.760)		(6.639)			(6.151)			(5.880)		(6.152)		(5.910)
Observations	517	517	517	517	517	517	510	477	510	477	517	517	517	517	517	517	504	453	504	453
R-squared	0.773	0.773	0.774	0.788	0.788	0.789	0.783	0.773	0.798	0.788	0.774	0.773	0.774	0.789	0.788	0.788	0.786	0.786	0.801	0.803
Mean Dep | Fight = 0	9.734	9.792	10.100	9.734	9.792	10.100	9.734	9.734	9.734	9.734	9.377	9.474	10.081	9.377	9.474	10.081	9.377	9.377	9.377	9.377
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes
Control	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed	Peace	Peace	Peace	Peace	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed	Peace	Peace	Peace	Peace
Including year after	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a city-year. Standard errors clustered at the city level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Sample is restricted to observations with at least one active gang. Fight is a dummy variable equal to one if a fight is recorded for that observation. Fight small is a dummy variable equal to one if a fight involving only small gangs is recorded for that observation. Fight large is a dummy variable equal to one if a fight involving at least one large gang is recorded for that observation. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. In Panels B and D, the outcome variable is the sum of arrests for drug and conspiracy causes. In columns 7,9,17,19 we exclude from the sample observations with fights involving large gangs. In columns 8,10,18,20 we exclude from the sample observations with fights involving small gangs. In columns 11 to 20 the fight dummies are equal to one also the period after the fight outbreak.

Table A7: Effects of fights, gang-year ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Panel A: number gang homicides																				
Fight	0.853***			0.862***							0.590***			0.676***						
	(0.113)			(0.104)							(0.0929)			(0.0900)						
Fight small		0.774***			0.784***		0.783***		0.795***			0.574***			0.665***		0.585***		0.663***	
		(0.110)			(0.103)		(0.110)		(0.104)			(0.0925)			(0.0884)		(0.0934)		(0.0905)	
Fight large			1.568***			1.497***		1.660***		1.649***			0.370			0.435		0.786		0.753
			(0.436)			(0.388)		(0.445)		(0.414)			(0.363)			(0.345)		(0.530)		(0.540)
Observations	1259	1259	1259	1259	1259	1259	1241	1013	1241	1013	1259	1259	1259	1259	1259	1259	1231	912	1231	912
R-squared	0.446	0.433	0.403	0.484	0.471	0.440	0.421	0.444	0.457	0.472	0.417	0.414	0.388	0.463	0.460	0.428	0.404	0.434	0.447	0.463
Mean Dep | Fight = 0	0.390	0.431	0.559	0.390	0.431	0.559	0.390	0.390	0.390	0.390	0.387	0.435	0.559	0.387	0.435	0.559	0.387	0.387	0.387	0.387
Panel B: number arrests																				
Fight	2.471*			0.975							0.856			0.334						
	(1.298)			(1.202)							(1.350)			(1.328)						
Fight small		1.150			-0.186		1.817		0.502			0.483			0.0860		0.568		0.381	
		(1.263)			(1.178)		(1.223)		(1.131)			(1.444)			(1.390)		(1.376)		(1.355)	
Fight large			17.50*			14.19		15.92		12.79			12.81**			9.983*		0.448		-1.034
			(10.24)			(9.686)		(10.16)		(9.575)			(5.988)			(5.405)		(4.252)		(4.349)
Observations	1259	1259	1259	1259	1259	1259	1241	1013	1241	1013	1259	1259	1259	1259	1259	1259	1231	912	1231	912
R-squared	0.832	0.831	0.834	0.854	0.854	0.856	0.826	0.856	0.852	0.874	0.831	0.831	0.833	0.854	0.854	0.855	0.817	0.871	0.844	0.888
Mean Dep | Fight = 0	33.729	35.178	33.690	33.729	35.178	33.690	33.729	33.729	33.729	33.729	33.627	35.631	33.278	33.627	35.631	33.278	33.627	33.627	33.627	33.627
Panel C: number gang homicides, no Naples																				
Fight	0.716***			0.675***							0.493***			0.543***						
	(0.152)			(0.150)							(0.112)			(0.105)						
Fight small		0.622***			0.580***		0.622***		0.572***			0.480***			0.547***		0.480***		0.534***	
		(0.136)			(0.131)		(0.136)		(0.132)			(0.118)			(0.113)		(0.118)		(0.116)	
Fight large			1.862***			1.834***		1.862***		1.841***			0.641**			0.554***		0.641**		0.563**
			(0.666)			(0.606)		(0.666)		(0.622)			(0.307)			(0.204)		(0.307)		(0.229)
Observations	684	684	684	684	684	684	679	616	679	616	684	684	684	684	684	684	675	578	675	578
R-squared	0.362	0.338	0.324	0.404	0.382	0.376	0.344	0.339	0.389	0.384	0.330	0.323	0.289	0.391	0.387	0.342	0.331	0.286	0.400	0.334
Mean Dep | Fight = 0	0.234	0.251	0.290	0.234	0.251	0.290	0.234	0.234	0.234	0.234	0.266	0.244	0.292	0.266	0.244	0.292	0.266	0.266	0.266	0.266
Panel D: number arrests, no Naples																				
Fight	-0.554			-0.918							1.070			0.445						
	(1.643)			(1.515)							(1.880)			(1.548)						
Fight small		-0.353			-0.571		-0.353		-0.553			0.938			0.241		0.938		0.323	
		(1.765)			(1.592)		(1.765)		(1.597)			(2.013)			(1.677)		(2.013)		(1.680)	
Fight large			-3			-5.043*		-3.000		-4.886			2.538			2.679		2.538		2.486
			(1.969)			(2.965)		(1.969)		(3.345)			(3.945)			(3.396)		(3.947)		(3.251)
Observations	684	684	684	684	684	684	679	616	679	616	684	684	684	684	684	684	675	578	675	578
R-squared	0.789	0.789	0.789	0.818	0.818	0.818	0.788	0.796	0.817	0.822	0.789	0.789	0.789	0.818	0.818	0.818	0.788	0.803	0.817	0.830
Mean Dep | Fight = 0	26.159	26.293	25.993	26.159	26.293	25.993	26.159	26.159	26.159	26.159	26.008	26.280	25.882	26.008	26.280	25.882	26.008	26.008	26.008	26.008
Gang FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	No	No	Yes	No	No	No	Yes	Yes	Yes	No	No	Yes	Yes
Control	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed	Peace	Peace	Peace	Peace	Mixed	Mixed	Mixed	Mixed	Mixed	Mixed	Peace	Peace	Peace	Peace
Including year after	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a gang-year. Standard errors clustered at the gang level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Sample is restricted to observations with at least one active gang. Fight is a dummy variable equal to one if a fight is recorded for that observation. Fight small is a dummy variable equal to one if a fight is recorded and the gang is small. Fight large is a dummy variable equal to one if a fight is recorded and the gang is large. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. In Panel B, the outcome variable is the sum of arrests for drug and conspiracy causes. In columns 7,9,17,19 we exclude from the sample observations with fights involving large gangs. In columns 8,10,18,20 we exclude from the sample observations with fights involving small gangs. In columns 11 to 20 the fight dummies are equal to one also the period after the fight outbreak.

Table A8: Fights and violence spillovers ← Return to text

Dep. Variable:	(1) Fight	(2) Fight	(3) Fight	(4) Fight	(5) N gang hom	(6) N gang hom	(7) N gang hom	(8) N gang hom	(9) N gang hom
Other fights	0.442***	0.417***	0.340***	0.330***	0.0925*				
	(0.0571)	(0.0584)	(0.0665)	(0.0680)	(0.0517)				
Other fights × small						0.0710		0.0762	0.0952
						(0.0629)		(0.0633)	(0.0867)
Other fights × large							0.117	0.123	0.192**
							(0.0863)	(0.0859)	(0.0873)
Observations	1412	1408	1408	1408	1408	1408	1408	1408	1239
R-squared	0.259	0.385	0.462	0.466	0.413	0.413	0.413	0.414	0.396
Location FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Band FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Control	All	All	All	All	All	All	All	All	Peace
Mean Dep | Other fights = 0	0.045	0.045	0.045	0.045	0.212	0.225	0.241	0.211	0.161

Notes: HDFE linear regression. Each observation is a band-location-year. Standard errors clustered at the band level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. Other fights is a dummy variable indicating whether the band is involved in other fights, in other locations, during the same period. This is missing for gangs present in only one location during the period. Small and large are dummy variables indicating whether the gang is small or large. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide.

Table A9: Fights and number of gangs ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: all gangs									
Fight	1.958***		1.396***	0.486***		0.356***	0.337***		0.339***
	(0.313)		(0.254)	(0.104)		(0.106)	(0.0984)		(0.0980)
Fight (lag)		1.796***	1.272***		0.144	0.125		-0.0313	-0.0475
		(0.310)	(0.224)		(0.112)	(0.115)		(0.0993)	(0.100)
Number gangs (lag)							0.320***	0.329***	0.325***
							(0.0401)	(0.0388)	(0.0380)
Observations	756	653	653	653	653	653	652	652	652
Mean Dep | Fight = 0	2.594	2.655	2.483	2.594	2.655	2.483	2.594	2.655	2.483
Panel B: small fights									
Fight	1.912***		1.350***	0.503***		0.376***	0.367***		0.372***
	(0.309)		(0.242)	(0.110)		(0.113)	(0.103)		(0.103)
Fight (lag)		1.728***	1.198***		0.147	0.115		-0.0335	-0.0639
		(0.311)	(0.225)		(0.110)	(0.113)		(0.0990)	(0.0996)
Number gangs (lag)							0.322***	0.329***	0.328***
							(0.0401)	(0.0392)	(0.0380)
Observations	756	653	653	653	653	653	652	652	652
Mean Dep | Fight = 0	2.617	2.683	2.532	2.617	2.683	2.532	2.617	2.683	2.532
Panel C: large fights									
Fight	2.631***		2.153***	0.382**		0.336*	0.345**		0.335*
	(0.704)		(0.584)	(0.164)		(0.172)	(0.165)		(0.174)
Fight (lag)		2.315***	1.655***		-0.0767	-0.0600		-0.211	-0.194
		(0.642)	(0.383)		(0.195)	(0.192)		(0.197)	(0.196)
Number gangs (lag)							0.326***	0.331***	0.331***
							(0.0399)	(0.0388)	(0.0391)
Observations	756	653	653	653	653	653	652	652	652
Mean Dep | Fight = 0	2.224	2.297	2.303	2.224	2.297	2.303	2.224	2.297	2.303
Location FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the number of active gangs in a location-year. Fight is a dummy variable indicating whether a fight has been recorded in that city-year. Panel A considers all fights. Panel B fights involving only small gangs. Panel C fights involving at least one large gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province.

Table A10: Fights and exits ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: all gangs						
Fight	0.209***		0.242***	0.211***		0.242***
	(0.0252)		(0.0280)	(0.0253)		(0.0280)
Fight (lag)		-0.0459**	-0.117***		-0.0466**	-0.117***
		(0.0207)	(0.0232)		(0.0207)	(0.0231)
Presence (lag)	0.890***	0.924***	0.904***	0.888***	0.922***	0.902***
	(0.0112)	(0.00712)	(0.00941)	(0.0112)	(0.00716)	(0.00944)
Observations	190575	190575	190575	190575	190575	190575
Panel B: small gangs						
Fight	0.228***		0.258***	0.229***		0.258***
	(0.0261)		(0.0288)	(0.0261)		(0.0288)
Fight (lag)		-0.0370	-0.111***		-0.0378	-0.110***
		(0.0247)	(0.0261)		(0.0247)	(0.0261)
Presence (lag)	0.844***	0.887***	0.861***	0.843***	0.886***	0.859***
	(0.0167)	(0.0118)	(0.0148)	(0.0166)	(0.0118)	(0.0147)
Observations	168553	168553	168553	168553	168553	168553
Panel C: large gangs						
Fight	0.0858**		0.0912**	0.0865**		0.0920**
	(0.0401)		(0.0436)	(0.0409)		(0.0439)
Fight (lag)		0.0181***	-0.0144		0.0155**	-0.0154
		(0.00648)	(0.0151)		(0.00687)	(0.0148)
Presence (lag)	0.975***	0.979***	0.975***	0.971***	0.976***	0.972***
	(0.00710)	(0.00664)	(0.00709)	(0.00756)	(0.00710)	(0.00756)
Observations	11011	11011	11011	11011	11011	11011
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Gang FE	No	No	No	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a location-gang-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is a dummy indicating whether the gang is recorded as present in the location-year. Panel A considers all gangs. Panel B only small gangs. Panel C only large gangs. Fight is a dummy variable indicating whether a fight has been recorded in that city-year. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province.

Table A11: Probability outcomes fights ← Return to text

Fighting outcome	(1) Win	(2) Draw	(3) Loose
Panel A: large			
Large	0.191***	-0.104**	-0.0868
	(0.0515)	(0.0477)	(0.0535)
Observations	403	403	403
Mean Dep | Large = 0	0.097	0.771	0.132
Panel B: scale			
Scale	0.0368***	-0.0264***	-0.0104
	(0.00644)	(0.00661)	(0.00857)
Observations	403	403	403
Mean Dep | Scale = 1	0.102	0.800	0.097
Panel C: large, num fights			
Large	0.158***	-0.0724	-0.0854
	(0.0542)	(0.0681)	(0.0722)
Num fights	0.0160	-0.0153	-0.000693
	(0.0201)	(0.0216)	(0.0192)
Observations	403	403	403
Mean Dep | Large = 0	0.097	0.771	0.132
Panel D: scale, num fights			
Scale	0.0334***	-0.0257*	-0.00777
	(0.0107)	(0.0151)	(0.0137)
Num fights	0.00722	-0.00158	-0.00564
	(0.0208)	(0.0250)	(0.0232)
Observations	403	403	403
Mean Dep | Scale = 1	0.102	0.800	0.097
Panel E: dependent variable is the number of fights			
Scale		0.468***	
		(0.106)	
Large			2.080**
			(0.880)
Observations	299	299	299
Mean Dep | Scale = 1	1.549	1.549	1.549

Notes: OLS linear regression. Each observation is a gang-year. Sample composed by observations with fights only. Standard errors clustered at the gang level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province. Outcomes are dummy variable indicating whether the gang has won, lost, or draw in the fight.

Table A12: Violence and seizures ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: log average quantity						
Number gang homicides	0.0199***	-0.0179***	-0.0171***			
	(0.00359)	(0.00614)	(0.00623)			
Fight				0.120*	0.0191	0.0178
				(0.0620)	(0.0720)	(0.0745)
Fight (lag)				0.136*	0.0231	0.0125
				(0.0695)	(0.0701)	(0.0691)
Observations	2208	2208	2208	1932	1932	1932
Mean Dep | Homicides/Fight = 0	0.147	0.147	0.147	0.144	0.144	0.144
Panel B: log average quantity, no outliers						
Number gang homicides	0.0166***	-0.0150**	-0.0146**			
	(0.00234)	(0.00737)	(0.00685)			
Fight				0.0695*	-0.00683	-0.00543
				(0.0388)	(0.0490)	(0.0512)
Fight (lag)				0.0864**	-0.00629	-0.0146
				(0.0422)	(0.0351)	(0.0365)
Observations	2208	2208	2208	1932	1932	1932
Mean Dep | Homicides/Fight = 0	0.114	0.114	0.114	0.111	0.111	0.111
Panel C: share seized quantity						
Number gang homicides	0.0197***	-0.0137***	-0.0140***			
	(0.000995)	(0.00200)	(0.00198)			
Fight				0.0504	-0.000807	-0.000846
				(0.0363)	(0.00801)	(0.00840)
Fight (lag)				0.0477	-0.00667*	-0.00692
				(0.0390)	(0.00383)	(0.00426)
Observations	2208	2208	2208	1932	1932	1932
Mean Dep | Homicides/Fight = 0	0.0056	0.0056	0.0056	0.0053	0.0053	0.0053
Panel D: share seized quantity, controlling for number of seizures						
Number gang homicides	-0.0113***	-0.0132***	-0.0134***			
	(0.00256)	(0.00174)	(0.00172)			
Fight				0.00418	-0.000938	-0.000878
				(0.00661)	(0.00807)	(0.00843)
Fight (lag)				-0.00405	-0.00837**	-0.00878*
				(0.00359)	(0.00415)	(0.00442)
Number of seizures	0.000744***	0.000391*	0.000396*	0.000588***	0.000476**	0.000492**
	(0.0000466)	(0.000220)	(0.000214)	(0.00000874)	(0.000238)	(0.000245)
Observations	2208	2208	2208	1932	1932	1932
Mean Dep | Homicides/Fight = 0	0.0056	0.0056	0.0056	0.0053	0.0053	0.0053
Panel E: quality						
Number gang homicides	0.00499	-0.229*	-0.0446			
	(0.0980)	(0.112)	(0.168)			
Fight				9.107**	16.35**	3.843
				(4.163)	(6.037)	(4.176)
Fight (lag)				-3.523	1.061	-2.752
				(4.057)	(9.240)	(4.808)
Observations	125	120	120	116	111	111
Mean Dep | Homicides/Fight = 0	31.278	31.278	31.278	28.616	28.616	28.616
City FE	No	Yes	Yes	No	Yes	Yes
Year FE	No	No	Yes	No	No	Yes
Drug FE	No	No	Yes	No	No	Yes

Notes: HDFE linear regression. Each observation is a city-drug-year. Standard errors clustered at the city level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable in Panels A and B is the logarithm of average quantity seized plus 1 in a given city-drug-year (total quantity divided by the number of seizures). Dependent variable in Panels C and D is the share of total quantity seized in a city, for a given drug and year. Fight is a dummy variable indicating whether a fight has been recorded in that city-year.

Table A13: Data and model ← Return to text

Data	Model	Source
Network matrix	$\mathbf{Y}_t\ \forall t$	Inteligence data - Neapolitan Police
Matrix of fights	$\mathbf{F}_t\ \forall t$	Inteligence data - Neapolitan Police
# of arrests for drug dealing and related crimes	Measure $O_{ty}\ \forall y, t$	Administrative data - DCSA
Geolicalized seized drug type, quantity and time	Measure $q_{yvt}\ \forall y, t, v$	Administrative data - DCSA
Geolocalized quality of drug	x_{ytv}	DCSA + Scientific Police
Average retail drug price over time and type	$p_{vt}\ \forall t, v$	Inteligence data - DCSA
Average wholesale drug price over time and type	$c_{vt}\ \forall t, v$	Inteligence data - DCSA
Distances between centroids of all markets	$g_{yy'}\ \forall y, y'$	Google maps API
Joint distribution of consumer groups and locations	$F(D_i, y_i)$	ISTAT + IPSAD survey

Table A14: Demand Estimation ← Return to text

	(1)	(2)	(3)	(4)	(5)
Panel A: Reduced form and Second Stage					
Num gang homicides	-0.00929***	-0.00965***	-0.0139**	-0.0113***	-0.0111***
	(0.00125)	(0.00117)	(0.00657)	(0.00116)	(0.00332)
Seizure share		0.585***		0.588***	0.588***
		(0.0429)		(0.0428)	(0.0434)
Selection fitted			0.00428		
			(0.00619)		
Num homicides					-0.000173
					(0.00227)
Observations	1584	1584	1149	1584	1584
R2	0.502	0.534	0.555	0.104	0.104
City FE	Yes	Yes	Yes	Yes	Yes
Drug Year FE	Yes	Yes	Yes	Yes	Yes
Selection	No	Yes	Yes	Yes	Yes
IV	No	No	No	Yes	Yes
KP-Stat				70.65	70.65
Mean Dep | Num homicides = 0	0.0072	0.0072	0.0071	0.0072	0.0072
First Stage					
Prison exit				-0.715***	-0.895***
				(0.085)	(0.153)
Panel B: Price - Second Stage					
Retail price				-4.18e-06	
				(9.29e-06)	
Observations				24	
KP-Stat				139.34	
First Stage					
Wholesale price				1.865***	
				(0.158)	
Panel C: Distance					
Distance				-0.0001252**	
				(0.0000555)	
Observations				71	
Panel D: Robustness					
		log-lin,+1	log-lin	log-log,+1	log-log
Num gang homicides		-0.0112962***	-0.0569754*	-0.1563478***	-0.7907693*
Elasticity		- 0.251	-0.030	-1.152	-0.116
Retail price		-4.18e-06	-0.0074266	-0.0003421	-0.3447*
Elasticity		-0.014	-0.319	-0.023	-0.345
Distance		-0.0001252**	-0.0193402***	-0.0267351***	-0.9722052***
Elasticity		-0.412	-0.440	-1.739	-0.929
Observations		1584	872	1584	872

Notes panel A: HDFE linear regression. Each observation is a city-year-drug. Sample is restricted to observations with at least one active gang. Standard errors clustered at the city level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the log of share of drug seized plus one. Seizures share is the share of the number of seizures plus one. Selection fitted is residual number of gang homicides from a first stage regression on the share of local police, additional details in Section 5. IV in columns (4) and (5) is the number of members of gangs' present in the area released from prison in the same period.
Notes panel B: OLS regression. Each observation is a drug-year. Standard errors clustered at the drug level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the drug-year fixed effect estimated in the previous panel, column (4). IV is the wholesale price.
Notes panel C: OLS regression. Each observation is a city. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the city fixed effect estimated in panel A. Independent variable is the average distance from the considered city to all the others, weighted by the consumer share, additional details in Section A.1.
Notes panel D: Each column replicates column (4) of panel A, for four different functional forms: (i) log-lin with dependent log of share of drug seized plus one–as in column (4), (ii) log-lin with dependent log of share of drug seized, (iii) log-log with dependent log of share of drug seized plus one and independent the log of the number of gang homicides plus one, (iv) log-log with dependent log of share of drug seized and independent the log of the number of gang homicides. For each model, the estimated coefficients are reported and the corresponding elasticity.

Table A15: Police Reaction Estimation ← Return to text

Share of arrests	(1) Local police	(2) DIA	(3) Conspiracy	(4) Drugs	(5) Local police	(6) DIA	(7) Conspiracy	(8) Drugs
Panel A: Police reaction								
Num gang homicides t	0.195**	-0.000141	-2.608***	-0.301	0.192**	-0.0000860	-2.583***	-0.345
	(0.0913)	(0.000114)	(0.866)	(0.691)	(0.0903)	(0.000113)	(0.860)	(0.643)
Num gang homicides $t-1$	-0.0161	-0.000119	3.319***	4.891***	-0.00544	-0.0000672	3.326***	4.751***
	(0.0338)	(0.000149)	(0.669)	(0.786)	(0.0375)	(0.000166)	(0.658)	(0.793)
Observations	442	644	644	644	442	644	644	644
R2	0.925	0.627	0.933	0.988	0.926	0.651	0.933	0.989
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes	Yes	Yes
Mean Dep | Num homicides = 0	0.3307	0.0022	0.4355	5.6624	0.3307	0.0022	0.4355	5.6624
Panel B: Police reaction - share								
Num gang homicides t	0.000463	0.00235***	-0.00260*	-0.000412	0.000485	0.00239***	-0.00257	-0.000433*
	(0.000765)	(0.000854)	(0.00155)	(0.000264)	(0.000774)	(0.000857)	(0.00156)	(0.000253)
Num gang homicides $t-1$	0.00360***	0.000686	0.00933***	-0.00233***	0.00365***	0.000712	0.00947***	-0.00237***
	(0.000511)	(0.000925)	(0.00127)	(0.000639)	(0.000530)	(0.000942)	(0.00127)	(0.000617)
Observations	442	644	644	644	442	644	644	644
R2	0.929	0.634	0.959	0.988	0.929	0.635	0.959	0.988
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes	Yes	Yes
Mean Dep | Num homicides = 0	0.0043	0.0087	0.0017	0.0045	0.0043	0.0087	0.0017	0.0045
Panel C: Police reaction - share, no Naples								
Num gang homicides t	-0.00888	0.000615	0.00640	0.000238	-0.00924	0.000642	0.00662	0.000247
	(0.00602)	(0.00167)	(0.0163)	(0.00164)	(0.00614)	(0.00170)	(0.0169)	(0.00164)
Num gang homicides $t-1$	0.0000586	0.00458**	0.0246**	0.00117	-0.0000697	0.00470**	0.0253**	0.00121
	(0.00324)	(0.00188)	(0.0102)	(0.00143)	(0.00353)	(0.00193)	(0.0106)	(0.00145)
Observations	437	637	637	637	437	637	637	637
R2	0.356	0.526	0.287	0.785	0.356	0.526	0.288	0.785
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes	Yes	Yes
Mean Dep | Num homicides = 0	0.0114	0.0097	0.0073	0.0089	0.0114	0.0097	0.0073	0.0089
Panel D: Police reaction - share, dummy								
$1\{$N. gang homicides $t>0 \mid t-1>0\}$	-0.00369***	-0.000435	0.00259**	-0.000195	-0.00916	-0.000939	0.00559*	-0.000422
	(0.00124)	(0.00125)	(0.00122)	(0.000438)	(0.00641)	(0.00253)	(0.00327)	(0.000824)
Observations	533	736	736	736	533	736	736	736
R2	0.917	0.633	0.952	0.987	0.917	0.633	0.952	0.987
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes	Yes	Yes
Mean Dep | Num homicides = 0	0.0049	0.0088	0.0017	0.0046	0.0049	0.0088	0.0017	0.0046
Panel E: Police reaction - share, dummy, no Naples								
$1\{$N. gang homicides $t>0 \mid t-1>0\}$	-0.00359	0.0000189	0.0103*	-0.000180	-0.00852*	0.0000414	0.0224**	-0.000394
	(0.00274)	(0.00149)	(0.00527)	(0.000882)	(0.00434)	(0.00261)	(0.0106)	(0.00144)
Observations	527	728	728	728	527	728	728	728
R2	0.354	0.529	0.250	0.783	0.355	0.529	0.255	0.783
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes	Yes	Yes
Mean Dep | Num homicides = 0	0.0121	0.0099	0.0066	0.0091	0.0121	0.0099	0.0066	0.0091

Panel F: Measures	(1)	(2)	(3)
Number of arrests (standardized)	DIA	Conspiracy	Drugs
Local police	0.0235***	0.0891***	0.0712***
	(0.00613)	(0.0103)	(0.00174)
Observations	533	533	533
R2	0.699	0.942	0.989
City FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes panels A-E: HDFE linear regression. Each observation is a city-year. Sample is restricted to observations with at least one active gang. Standard errors clustered at the city level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the number of arrests, panel A, or the share of arrests, panel B to E, from local police–columns (1) and (5), reported in the DIA reports–columns (2) and (6), for conspiracy–columns (3) and (7), and for drug related crimes–columns (4) and (8). In panels A to C the independent variable is the number of gang homicides in the period, and the period before. In panels D and E is a dummy variable indicating whether one of the two variables is strictly positive.

Notes panel F: HDFE linear regression. Each observation is a city-year. Sample is restricted to observations with at least one active gang. Standard errors clustered at the city level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the share of arrests by the local police. Independent variable is the share of arrests reported in the DIA reports in column (1), the share of arrests for conspiracy in column (2), and the share of arrests for drug related crimes in column (3).

Table A16: Violence generation ← Return to text

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: number gang homicides						
Fight small	0.194		0.262	0.248		0.319**
	(0.231)		(0.162)	(0.190)		(0.156)
Fight large		1.725**	1.768**		0.601	1.680**
		(0.760)	(0.788)		(0.407)	(0.718)
Observations	525	525	525	212	54	258
R2	0.706	0.717	0.718	0.716	0.765	0.722
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	Ever small war	Ever large war	Ever war
Mean Dep | Fight = 0	0.282	0.164	0.097	0.635	0.585	0.153
Panel B: number gang homicides, no Naples						
Fight small	0.389***		0.383***	0.367**		0.399***
	(0.130)		(0.129)	(0.137)		(0.135)
Fight large		1.050*	1.039*		0.830	1.051*
		(0.579)	(0.575)		(0.567)	(0.566)
Observations	517	517	517	204	46	250
R2	0.271	0.288	0.325	0.326	0.494	0.338
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	All	All	Ever small war	Ever large war	Ever war
Mean Dep | Fight = 0	0.112	0.128	0.097	0.152	0.154	0.153
Panel C: number gang homicides - spillovers						
Other fights small		0.0762	0.0952		0.0762	-0.0219
		(0.0633)	(0.0867)		(0.0715)	(0.0997)
Other fights large		0.123	0.192**		0.0646	0.0681
		(0.0859)	(0.0873)		(0.0943)	(0.0411)
Observations		1408	1239		808	753
R2		0.414	0.396		0.259	0.232
Band FE		Yes	Yes		Yes	Yes
Location FE		Yes	Yes		Yes	Yes
Year FE		Yes	Yes		Yes	Yes
Naples		Yes	Yes		No	No
Control		All	Peace		All	Peace
Mean Dep | Other fights = 0		0.212	0.161		0.112	0.089

Notes panels A - B: HDFE linear regression. Each observation is a city-year. Sample is restricted to observations with at least one active gang. Standard errors clustered at the city level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the number of gang homicides. Fight small is a dummy variable equal to one if a fight involving only small gangs is recorded for that observation. Fight large is a dummy variable equal to one if a fight involving at least one large gang is recorded for that observation. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province. In column (4) we restrict the control group to cities where fight small is ever equal to one during the period, in column (5) the same for fight large, and column (6) for both. In panel B we exclude the city of Naples from the sample.

Notes panel C: HDFE linear regression. Each observation is a location-band-year. Sample is restricted to observations with at least one active gang. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Dependent variable is the number of gang homicides. Other fight small is a dummy variable equal to one if the gang present in the location is involved in a fight in another location in the same period, and the gang is small. Other fight large is a dummy variable equal to one if the gang present in the location is involved in a fight in another location in the same period, and the gang is large. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province. In columns (3) and (6) we restrict the sample to location with no fight recorded in the period.

Table A17: Violence generation - robustness ← Return to text

	Unit level				City level				City level, no Naples			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Fight - scale = 1	0.226	0.262	0.227	0.264	0.300	0.294	0.334	0.369	0.313	0.316	0.313	0.316
	(0.190)	(0.210)	(0.190)	(0.210)	(0.305)	(0.375)	(0.320)	(0.409)	(0.281)	(0.326)	(0.281)	(0.326)
Fight - scale = 2	0.900***	0.899***	0.899***	0.899***	0.219	0.237	0.185	0.253	0.484*	0.510**	0.484*	0.510**
	(0.197)	(0.186)	(0.195)	(0.184)	(0.345)	(0.352)	(0.336)	(0.303)	(0.249)	(0.245)	(0.249)	(0.245)
Fight - scale = 3	0.230**	0.230*	0.232**	0.234*	0.503**	0.504*	0.478*	0.543	0.316**	0.329*	0.316**	0.329*
	(0.111)	(0.124)	(0.111)	(0.123)	(0.249)	(0.250)	(0.281)	(0.333)	(0.154)	(0.162)	(0.154)	(0.162)
Fight - scale = 4	0.830***	0.791**	0.809**	0.765**	1.199**	1.189**	1.115**	0.957	1.193**	1.204**	1.193**	1.204**
	(0.314)	(0.300)	(0.308)	(0.295)	(0.575)	(0.572)	(0.550)	(0.589)	(0.573)	(0.562)	(0.573)	(0.562)
Fight - scale = 5, >5	0.791	0.684	0.434	0.467	15.30***	15.23***	3.269*	3.320*			0.0192	0.0557
	(0.566)	(0.577)	(0.309)	(0.309)	(0.279)	(0.252)	(1.690)	(1.671)			(0.0967)	(0.120)
Fight - scale = 7	0.669	0.737			1.042	1.066			0.0192	0.0557		
	(0.596)	(0.587)			(0.812)	(0.778)			(0.0967)	(0.120)		
Fight - scale = 8	-0.0471	0.00189			-1.698***	-1.768***						
	(0.206)	(0.213)			(0.281)	(0.260)						
Fight - scale = 9	0.502	0.586			-1.058***	-1.064***						
	(0.650)	(0.662)			(0.288)	(0.271)						
Observations	753	408	753	408	525	258	525	258	517	250	517	250
R2	0.434	0.442	0.429	0.437	0.920	0.926	0.723	0.727	0.336	0.351	0.336	0.351
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	All	Ever war	All	Ever war	All	Ever war	All	Ever war	All	Ever war	All	Ever war
Winsorized	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Mean Dep | Fight = 0	0.111	0.166	0.111	0.166	0.097	0.153	0.097	0.153	0.097	0.153	0.097	0.153

Notes: HDFE linear regression. Each observation is a city-year in panel A and location-year in panel B. Sample is restricted to observations with at least one active gang. Standard errors clustered at the city level in panel A and location level in panel B. ***, **, * indicate significance at the 1, 5, and 10% levels, respectively. Dependent variable is the number of gang homicides. Fight-scale $= x$ is a dummy equal to one if a fight involving gangs of at most scale x is recorded for that observation. The scale of a gang, in a given year, is the number of locations in which it is present. Locations are districts within Naples and cities in the province. In columns (3) and (4) fights involving gangs with maximum scale > 5 are included in the dummy for 5. In columns (2) and (4) we restrict the control group to cities/locations where a fight is ever observed during the sample period.

Table A18: Model fit - quantity ← Return to text

Seized	(1)	(2)	(3)	(4)
Simulated	0.665***	0.665***	0.665***	3.552***
	(0.0188)	(0.0188)	(0.0188)	(0.615)
Observations	1584	1584	1584	1584
R2	0.442	0.442	0.442	0.490
Drug FE	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes
City FE	No	No	No	Yes

Notes: HDFE local linear regression. Dependent variable is log share of drug seized in a given city, year, and for a given drug. Independent variable is the log share simulated by the model.

Table A19: Model fit - fights, maximum scale ← Return to text

	(1) Data	(2) Model	(3) Random	(4) Fixed	(5) Data	(6) Model	(7) Random	(8) Fixed	(9) Data	(10) Model	(11) Random	(12) Fixed
Maximum scale	-0.0769***	-0.0354	0.0130	0.00188	-0.0865***	-0.0650	0.0377	0.0401*	-0.0674**	0.00166	0.0412	0.0369
	(0.0194)	(0.0302)	(0.0136)	(0.00807)	(0.0296)	(0.0573)	(0.0445)	(0.0210)	(0.0323)	(0.0538)	(0.0461)	(0.0230)
Number gangs	0.179***	0.156***	0.157***	0.417***	0.294***	0.0882	0.169***	0.446***	0.289***	0.123*	0.158**	0.448***
	(0.0255)	(0.0330)	(0.0158)	(0.0176)	(0.0605)	(0.0734)	(0.0595)	(0.0269)	(0.0593)	(0.0683)	(0.0621)	(0.0278)
Observations	756	756	756	756	753	753	753	753	753	753	753	753
R2	0.150	0.104	0.111	0.762	0.365	0.600	0.214	0.806	0.382	0.655	0.218	0.808
P-value scale		0.229	0.000	0.000		0.726	0.028	0.000		0.241	0.066	0.005
P-value number		0.520	0.405	0.000		0.011	0.160	0.020		0.024	0.146	0.014
Location FE	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. Maximum scale is the maximum scale of gangs active in that location in that period. Columns (1), (5), (9) have as dependent variable a dummy indicating a fight in the data. Columns (2), (6), (10) have as dependent variable the sum of fighting probabilities predicted by the model. Columns (3), (7), (11) a dummy indicating a randomly predicted fight. Columns (4),(8),(12) the sum of randomly predicted fighting probabilities, randomly drawn from a normal distribution with the same moments as the predicted fighting probabilities. Dependent variables are standardized in all models. P-values indicate the equality of coefficients with the one for columns (1), (5), (9).

Table A20: Model fit - fights, additional covariates ← Return to text

	(1) Data	(2) Model	(3) Random	(4) Fixed
Panel A: Average scale				
Average scale	-0.0980***	-0.157***	0.00485	-0.0165
	(0.0269)	(0.0444)	(0.0188)	(0.0103)
Number gangs	0.147***	0.109***	0.155***	0.415***
	(0.0241)	(0.0359)	(0.0170)	(0.0178)
Share consumers	0.231	0.376	0.00699	0.0538
	(0.191)	(0.243)	(0.132)	(0.0900)
Average distance	-0.000917	-0.0156	-0.00753*	0.00000784
	(0.00479)	(0.0148)	(0.00411)	(0.00251)
Observations	756	756	756	756
R2	0.143	0.166	0.113	0.763
P-value scale		0.236	0.000	0.005
P-value number		0.334	0.771	0.000
P-value consumers		0.631	0.333	0.411
P-value distance		0.276	0.312	0.848
Panel B: Maximum scale				
Maximum scale	-0.0824***	-0.0451	0.0113	0.00118
	(0.0202)	(0.0299)	(0.0140)	(0.00811)
Number gangs	0.184***	0.152***	0.152***	0.418***
	(0.0261)	(0.0347)	(0.0172)	(0.0180)
Share consumers	0.249	0.290	-0.00339	0.0370
	(0.194)	(0.253)	(0.132)	(0.0899)
Average distance	-0.00397	-0.0182	-0.00719*	-0.000105
	(0.00467)	(0.0154)	(0.00407)	(0.00253)
Observations	756	756	756	756
R2	0.155	0.126	0.114	0.978
P-value scale		0.288	0.000	0.000
P-value number		0.420	0.276	0.000
P-value consumers		0.892	0.290	0.325
P-value distance		0.311	0.609	0.423
Location FE	No	No	No	No
Year FE	No	No	No	No

Notes panel A: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. Average scale is the average scale of gangs active in that location in that period. Column (1) has as dependent variable a dummy indicating a fight in the data. Share consumers is the share of estimated consumers in the location (please refer to section D for additional details). Average distance is the average distance in km from the location to all the others. Column (1) has as dependent variable the sum of fighting probabilities predicted by the model. Column (3) a dummy indicating a randomly predicted fight. Column (4) the sum of randomly predicted fighting probabilities, randomly drawn from a normal distribution with the same moments as the predicted fighting probabilities. Dependent variables are standardized in all models. P-values indicate the equality of coefficients with the one for column (1).

D. OTHER DATA

Chemical analysis For a subsample of the seizures dataset, information is also available on the chemical composition of the seized drugs, namely the percentage of pure substance in each sample. When a penal proceeding follows a seizure–i.e., when the defendant does not plead guilty immediately–the judge is required to request a chemical analysis from the forensic police as a preliminary step. Italian law establishes a minimum percentage of pure substance for the crime of drug dealing. The forensic police perform the analysis, record the results in an official folder, and transmit the folder to the judge. Two anonymized random samples of these forensic analyses are constructed. First, 567 analyses are drawn at random from all relevant folders in the province of Naples for the years 2015–2022. Each folder reports the defendant's identity and the percentage of pure substance. Second, these analyses are merged internally with a confidential version of the drug-seizure dataset containing identifiers for the recipients of the seizures. The merged dataset is then anonymized and released to the research team. Because seizures are highly concentrated in the city of Naples, 86 percent of the chemical analyses in this purely random sample originate from the provincial capital. To increase geographic coverage, a second sample of 660 analyses is drawn by stratifying on city–year–drug cells and selecting a small number of observations from each cell so as to maximize the number of strata represented. For every analysis in the two samples, seizure characteristics (city, year, quantity) and the percentage of pure substance are observed. For each city–year–drug cell, the within-cell dispersion and mean level of drug purity are computed.

Average prices Each semester the Italian antidrug authority requires the three main enforcement agencies–*Polizia di Stato*, *Arma dei Carabinieri*, and *Guardia di Finanza*–to submit an official form reporting the average retail and wholesale prices, average purity, and average quantity for each type of drug. All available forms are collected and used to compute, for each year and drug type, the average retail and wholesale prices.[38]

Homicides Violence is measured using intelligence records of homicides and attempted homicides in the province of Naples for the period 2015–2022. For each event, the dataset reports the victim, location, and date. Using these variables, gang-related episodes are identified by matching each event with online

[38]The wholesale price of heroin is missing for 2019–2021 and is imputed as the average of the values observed in 2018 and 2022.

gang mentions.

Arrests Data on arrests are obtained using administrative records of the antidrug autority. For each city and year in the sample, the dataset reports the number of arrests for drug-related crimes and conspiracy.

Population Population size and demographics are measured using administrative data from ISTAT. For each city and year, the dataset reports the number of residents by sex (male or female) and age group (0–9, 9–24, 25–49, 50–64, >64). From these classifications, four demographic groups are defined as the intersections of two dummy variables: (i) male and (ii) young (age below 49). The number of residents in each demographic group is then computed for every city. For the city of Naples, district-level projections are obtained using the 2011 census. The census provides, for each district, the percentage of residents by sex and age. The estimated shares are then multiplied by the total number of residents in each demographic group to obtain district-level population counts.

Distances Distances between units are measured using the Google Maps API. For each pair of city or district centroids, the average car travel time–net of traffic conditions–is computed along with the corresponding distance in kilometers. In the analysis, travel time is used as the distance measure.

Consumption survey The propensity to consume different types of drugs across demographic groups is estimated using aggregate national statistics from the IPSAD survey. This Italian survey, conducted under the guidelines of the European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), satisfies the Council of Europe's requirement for the first of its five epidemiological indicators. The sample, drawn from civil registry lists, covers individuals aged 15–84 and, since 2005, has included about 85,000 respondents, allowing prevalence estimates at the regional level. The publicly available data report average consumption propensities for different drugs across demographic groups.

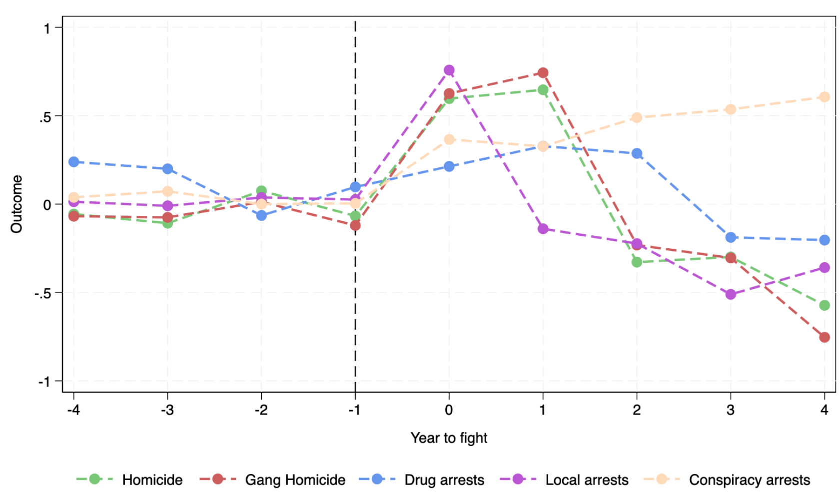
City survey City characteristics and municipal police activity are measured using data from the OpenCivitas city survey. OpenCivitas is a transparency portal promoted by the Ministry of Economy and Finance and SOSE that provides information on local authorities. Each year–except in 2020, when the survey was suspended due to the COVID-19 pandemic–cities in Italy are asked to report balance-sheet values and details on the services they provide. Of particular relevance for this study, the survey records the resources allocated to the municipal police and key outcomes of their activity, such as the number of arrests.

E. additional empirics

E.1. Data quality

This paper draws on the intersection of administrative and intelligence data. Assuming administrative data are complete and accurate, the quality of the intelligence data is assessed by examining how administrative measures vary with changes in intelligence outcomes. Specifically, I test whether administrative data on violence, arrests, and drug seizures respond as expected to changes in intelligence measures of presence and fights. Consistent and precise responses provide evidence of the reliability of the intelligence data.

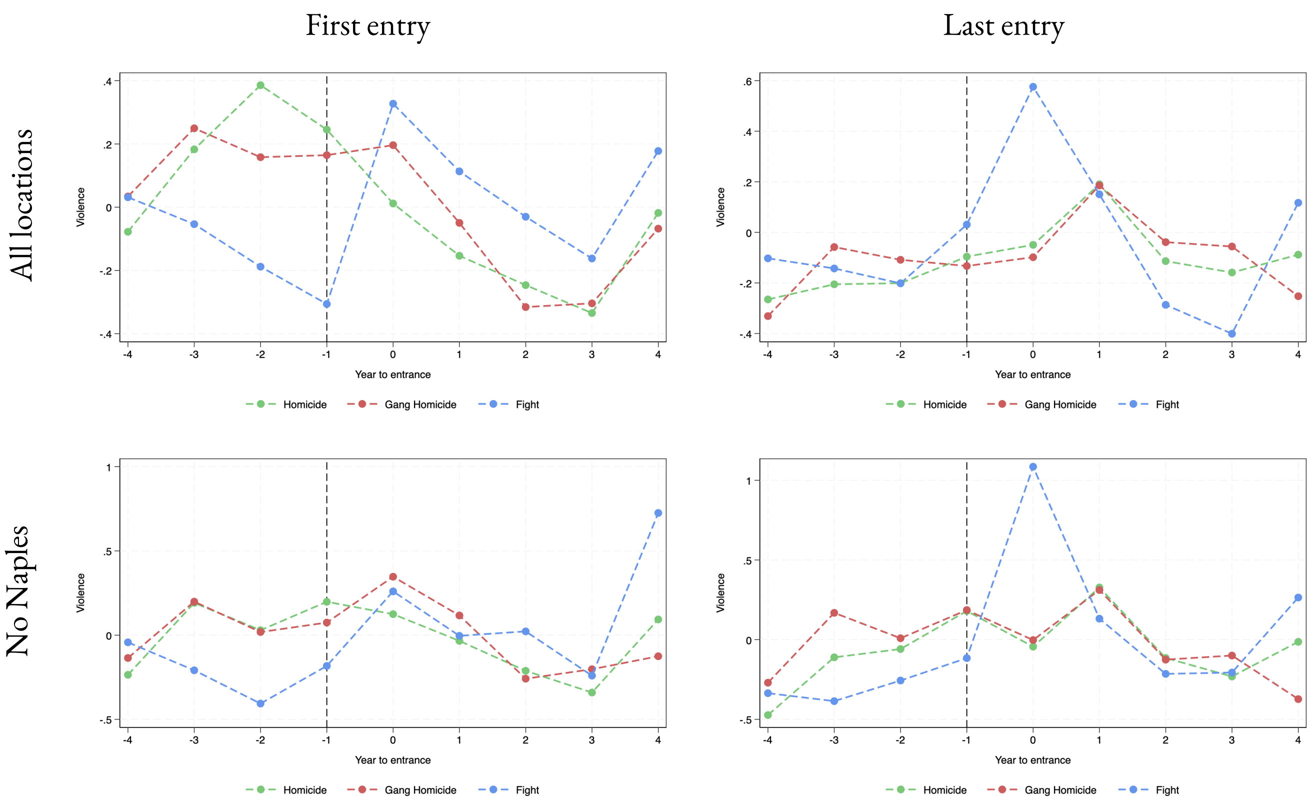
As a first step, I examines how violence and arrests evolve around the onset of fights. If fights are correctly recorded, violence and enforcement activity should rise when fights occur, without significant increases in the preceding period and following a plausible temporal pattern. Figure A26 plots standardized counts of homicides, gang homicides, drug arrests, municipal-police arrests, and arrests for conspiracy against time relative to the first fight in each area, controlling for city fixed effects. All outcomes remain close to the average before the outbreak, with no sign of shocks or pronounced pre-trends. At the onset of fights, all outcomes display sharp increases. Such clear temporal concentration would be unlikely in the presence of substantial measurement error, supporting the accuracy of the intelligence data.



Notes: The figure shows the standardized average of different outcomes relative to the first fight recorded in a location. Each observation is a location-year. Sample is restricted to observations with at least one active gang. Locations are districts within Naples and cities in the province. City fixed effects are included.

Figure A26: Data quality - variation with fights

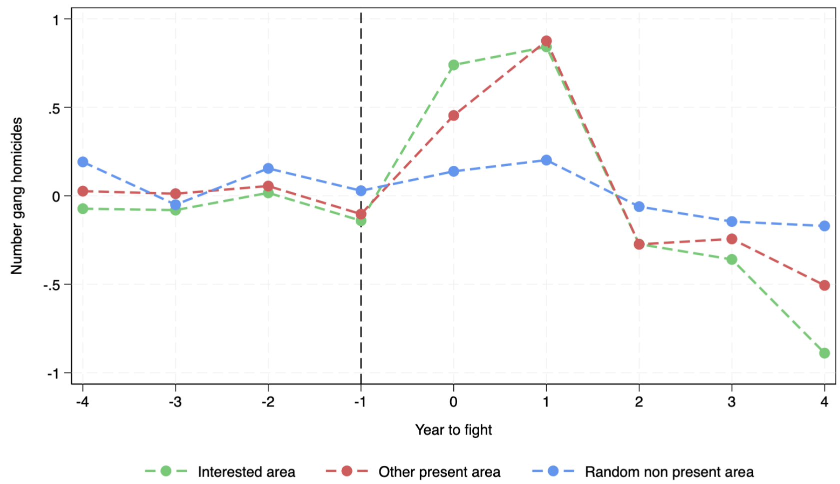
The next step is to examine presence data. The entry of a new gang into an area is expected to raise violence, as highlighted in the literature (Bruhn et al., 2021). Figures A27 illustrates this mechanism by plotting standardized counts of homicides and gang homicides, together with the probability of fights, against time relative to the first and last recorded gang entry in each area, including and not including Naples city in the sample. Although not always cleanly, at the time of entry, the probability of fights rises sharply and violence increases in the subsequent period.



Notes: The figure shows the standardized average of different outcomes relative to the first/last entry recorded in a location, considering all locations and excluding Naples city from the sample. Each observation is a location-year. Sample is restricted to observations with at least one active gang. Locations are districts within Naples and cities in the province. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. City fixed effects are included.

Figure A27: Data quality - variation with entry

To conclude, the interaction between the two intelligence datasets is examined. When a gang operates in multiple locations, violence is expected to rise with fights, to some extent, in some/all those locations. By contrast, violence should not increase in locations where the fighting gangs are absent, providing a placebo test. Figure A28 shows standardized gang homicides against time relative to the first fight, controlling for location fixed effects. Three types of areas are considered: (i) the location where the fight occurs, (ii) other locations where one of the fighting gangs operates, and (iii) randomly selected locations where neither gang is present. At the outbreak of fights, gang homicides increase in the first two location types with comparable magnitude, although the rise is smaller in the other affected areas during the initial period, as expected. No change is detected in the placebo locations. In the presence of substantial measurement error in the presence data–either in timing or in the recording of gang locations–these clear temporal patterns and differentiated spillover effects would be unlikely.



Notes: The figure shows the standardized average of the number of gang homicides relative to the first fight of a gang. Each observation is a location-gang-year. Three different averages are plotted: the average number of homicides in the location interested by the fight, the average of all the other locations where the gang was present but were not interested by the fight, a random location where the gang was not present. Locations are districts within Naples and cities in the province. City fixed effects are included.

Figure A28: Data quality - variation with presence

These exercises show that administrative measures respond in a reasonable and precise manner to changes in the intelligence data, supporting their reliability. This evidence is not, however, a formal test of data quality.

E.2. Additional descriptive evidence

Correlations Figure A29 shows the average probability of a fight, along with 95% confidence intervals, for locations with no gangs and for quintiles of the distribution of the number of gangs present (0–20, 20–40, 40–60, 60–80, 80–100). The probability is zero in locations without active gangs, rises to about 10% in locations with a single gang, and then increases roughly linearly with the number of gangs, reaching 35% in the top quintile. In short, the likelihood of conflict increases with the number of active gangs. Figures A30 and A31 replicate the analysis for homicides and gang-related homicides, respectively. Figure A32 reports results for non-gang homicides as a placebo exercise and finds no correlation.

Figure A33 shows the average probability of a fight, with 95% confidence intervals, for locations with no gangs and for quintiles of the distribution of average gang scale. As in previous graph, there are no fights in locations without active gangs. Where only very small gangs are present (Q1), the average fight probability is high, around 30%. As the average scale of active gangs increases, this probability declines, reaching about 4% in the fifth quintile. Thus, the likelihood of conflict is 26 percentage points higher in locations dominated by small gangs than in those dominated by large gangs. The same pattern emerges when considering homicides (Figure A34) and gang-related homicides (Figure A35). Figure A36 presents the placebo using non-gang homicides and shows no relationship. Figure A37 replicates the analysis using the maximum, rather than the average, gang scale in the location. Figure A38 reports the average fight probability by gang scale, rather the average scale in the location.

Figure A39 shows how the probability of gang homicides varies with the standard deviation of gang size. An inverse U-shaped relationship emerges. When active gangs are very similar in size—whether uniformly small or large—the probability of gang homicide is relatively low. As differences in size grow, violence rises, reaching a maximum when the standard deviation of gang size is around one. Beyond this point, as inequality in gang size becomes pronounced, the probability of gang homicide declines to its lowest level. Violence is therefore limited both when power is balanced, making conflict outcomes too unpredictable, and when it is highly unbalanced, making outcomes too predictable. Figure A40 shows a similar pattern for fights.

Gang entry Here I focus on changes in number of gangs in the average scale determined by gang entry. Entry affects both margins: it increases the number of active gangs and changes average scale.

Figure A41 reports an event–study analysis of three cases: the last entry of any gang in a location (green), the last entry of a small gang (scale ≤ 3, red), and the last entry of a large gang (scale > 3, blue). The entry of a new gang is associated with a 30% increase in fight probability in the first period and an 18% increase in the second period relative to the omitted category, consistent with a positive relationship between violence and the number of competitors (Bruhn et al., 2021). The effects are similar for the entry of a small gang, which also reduces average gang scale. By contrast, the entry of a large gang is associated with only a small and statistically insignificant 7% increase in the first period, followed by a statistically significant 10% decline in the second. Thus, the entry of a large gang seems not raise the likelihood of fighting and may even reduce it. This pattern aligns with the descriptive evidence: while large-gang entry increases the number of competitors and may stimulate conflict, it also raises average gang scale, which is negatively associated with fighting probability.

The absence of pre-trends is reassuring about the strength of the parallel trend assumption. Figure A42 presents the "raw data" version of the event study, plotting average fight probabilities in the years surrounding the entry of a gang, a small gang, and a large gang. The patterns closely match the event–study results. Figure A43 shows the distribution of entry years used in estimation, indicating that entries of both small and large gangs are widely dispersed across the sample. Figure A44 replicates this using the first, rather than the last, entry in each location. Figures A45 and A46 replicate the event–study and raw–data graphs for the probability of gang homicide, while Figures A47 and A48 do so for the number of gang homicides. Figures A49 and A50 replicate the analysis for first, rather than last, entry.

Table A21 reports results for all gang entries, not only the first or last. In Panel A, I regress a fight indicator on dummies equal to one in the period when a new gang (small or large) enters a location. The results mirror the event–study evidence: small gang entry increases fight probability, while large gang entry has no significant effect. Panels B and C replicate the analysis for the probability and number of gang homicides. Columns 3–4 add location fixed effects, and the results remain consistent. Columns 5–6 add year fixed effects with similar conclusions. The event–study graphs (Figures A41, A45, A47) suggest dynamic effects extending beyond the entry period. To account for this, Panels F–H use a dummy equal to one in both the entry period and the subsequent year. Across all specifications, small gang entry is associated with increased probability of fights during and after entry, while large gang entry shows either no effect or a decline. As a validation exercise, Panels D–E examine entry effects on average gang scale and the number of active gangs during the entry year, while Panels I–L extend the analysis to include the year after. As expected, small gang entry reduces average scale, whereas large gang entry increases it, confirming effects on the second channel. As for the first margin, any gang entry increases the number of active gangs, with larger effects for small gangs, though not statistically different.

Table A22 replicates Table A21, Panels A, B, C, F, G, and H, using the minimum gang scale over the study period as the scale measure. Table A23 replicates Table A21, Panels A and B, defining large gangs as those with scale above the average in the location (rather than above three). Panels C and D apply the same logic using the minimum scale over the period. Table A24 further examines sensitivity to the scale measure by replicating with four alternative definitions: (i) the benchmark (number of locations where the gang is present), (ii) the population living in those locations, standardized, (iii) the number of locations weighted by estimated consumer shares for cocaine, heroin, and marijuana, and (iv) the number of locations weighted by the share of seized drug quantity. Results are consistent across almost all definitions. Figures A51–A54 replicate the event–study graphs using alternative definitions of large gangs. Specifically, I estimate models separately for the entry of scale–1 gangs (present in only one location), and for gangs with scale greater than 1, 2, 3 (benchmark), and 4. As the threshold for defining large gangs increases, the attenuating effect on fight probability becomes stronger. Figures A55–A58 replicate the analysis for first rather than last entry. Table A25 reports the corresponding fixed-effects regressions, replicating Table A21, Panel A, columns 5–6, for these alternative definitions.

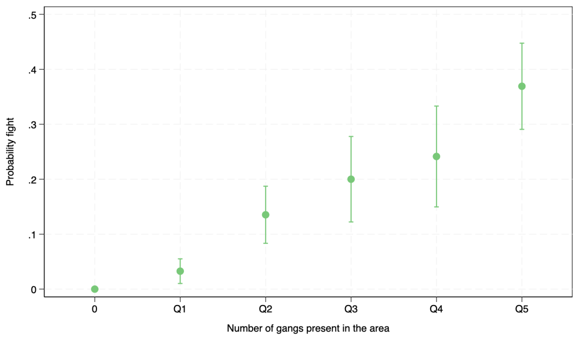
To conclude, Table A26 examines how fight probabilities change when an already active gang enters *another* location. Panel A replicates Table A2, Panel A, for comparison. Panel B presents the reduced form results, estimated without fixed effects (columns 1–2), with location fixed effects (columns 3–4), and with both location and year fixed effects (columns 5–6), controlling or not for the number of active gangs in the area. With fixed effects included, the entry of an active gang into another location–an event that mechanically raises the average gang scale without directly affecting the location itself–is associated with a large reduction in fight probability (–20%). The coefficient is imprecisely estimated but quantitatively consistent with the previous evidence. Panels C and D use entry into another location as an instrument for average gang scale. Panel C reports the first stage, regressing average scale on other-location entry. The coefficients are positive and statistically significant across all specifications. Panel D reports the second stage, with Kleibergen–Paap statistics around 26 in all fixed-effects models. The second-stage coefficient is negative and relatively close in magnitude to the OLS estimate (in column 5, a one-unit increase in average scale reduces fight probability by 15% in OLS and 28% in IV), though again estimated imprecisely.

Covid-19 Figure A59 shows the number of active gangs in the province (blue) and the average number of gangs per area (red). The total number of gangs rises from about 75 in 2015 to roughly 85 in 2022, with a sharp increase during the COVID-19 period, continuing until the end of the restrictions in 2021. The average number of gangs per area remains above one throughout, indicating the presence of multiple gangs within each area. As the total number of gangs grows, the average per area rises accordingly. Figure A60 reports the time trend of violence. Homicides remain relatively stable, with a slight decline over the sample period, while attempted homicides display a pronounced spike during the COVID-19 period.

Externalities Arrest data from DIA reports (e.g., DIA, 2020) allow measurement of arrest spillovers across gangs, as these records include the gang affiliation of arrested individuals–information not available in other arrest datasets. These results must be interpreted cautiously because DIA-reported arrests represent a selected subset of total arrests. Figures A61 and A62 plot the average number of DIA-reported arrests for each gang around (i) its first fight (green) and (ii) the first fight in an area where the gang was present but not directly involved (yellow). Results are shown for both raw counts and deviations from gang means. In both specifications, no pre-trends are observed before fights, and DIA-recorded arrests increase sharply at the outbreak of fights. For gangs present in the area but not directly involved, different dynamics emerge. Arrests rise slightly in the period just before the fight. At the time of the fight and immediately afterward, arrest rates are similar to pre-fight levels. Starting in the second year after the fight, however, arrests increase substantially for these non-involved gangs. Formal event-study regressions without fixed effects (Figure A63) and with fixed effects (Figure A64) confirm these patterns.[39]

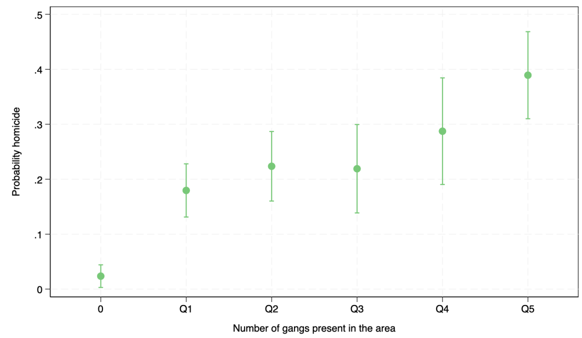
Drug quality Figure A65 displays the distribution of drug quality, measured as the percentage of pure substance, for the three drug types considered in this study. Average purity is 62 percent for cocaine, 22 percent for heroin, and 14 percent for marijuana. Substantial within-drug variation is also observed: for example, cocaine samples range from 2 percent to 98 percent purity, indicating coverage of the full spectrum of drug quality.

[39]Period −2 is used as the reference group to account for the pre-fight increase seen in the raw data.



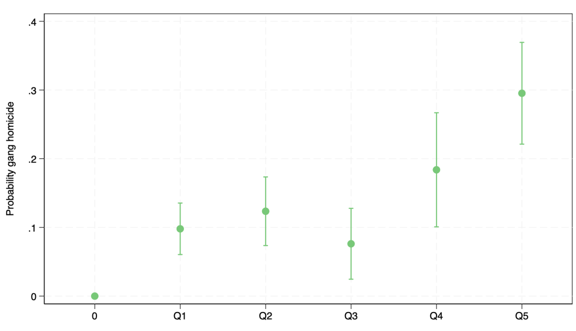
Notes: This figure presents the average fight probability for locations-years with no gangs, and for the five quintiles of the number of active gangs. Each observation is a location-year. Location is a district within Naples, and a city in the province of Naples. 95% confidence interval shown.

Figure A29: Prob. fight by number of gangs



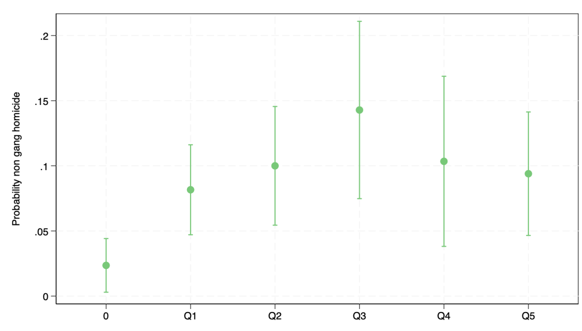
Notes: The figure shows the average probability of homicides for location-years with no gangs and for five quintiles of the number of active gangs. Each observation is a location-year. A location is a district within Naples, and a city in the province. 95% confidence intervals shown.

Figure A30: Prob. homicide by number of gangs



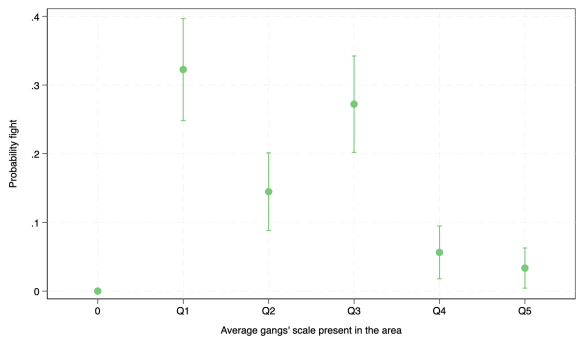
Notes: The figure shows the average probability of gang homicides for location-years with no gangs and for five quintiles of the number of active gangs. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. Each observation is a location-year. Locations are districts within Naples and cities in the province. 95% confidence intervals shown.

Figure A31: Prob. gang homicide by number of gangs



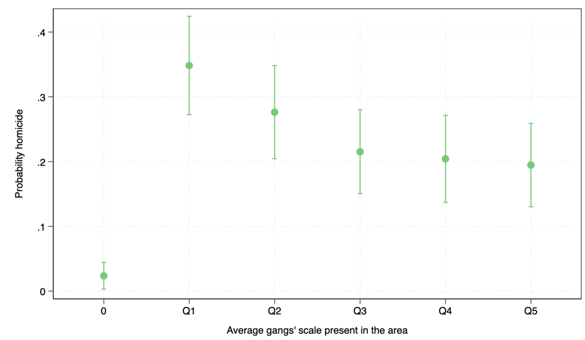
Notes: The figure shows the average probability of non-gang homicides for location-years with no gangs and for five quintiles of the number of active gangs. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. Each observation is a location-year. Locations are districts within Naples and cities in the province. 95% confidence intervals shown.

Figure A32: Prob. non-gang homicide by number of gangs



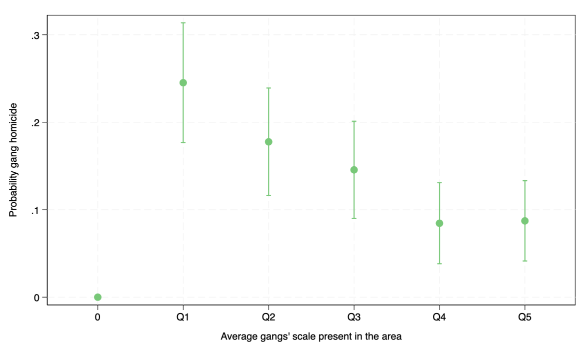
Notes: This figure presents the average fight probability for locations-years with no gangs, and for the five quintiles of the average gang scale. The scale of a gang, in a given year, is the number of locations in which it is present. Each observation is a location-year. Location is a district within Naples, and a city in the province of Naples. 95% confidence interval shown.

Figure A33: Prob. fight by average gang scale



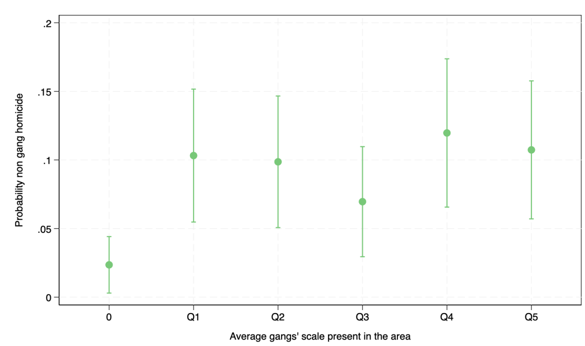
Notes: The figure shows the average probability of homicides for location-years with no gangs and for five quintiles of the average scale of gangs present in the area. The scale of a gang, in a given year, is the number of locations in which it is present. Each observation is a location-year. Locations are districts within Naples and cities in the province. 95% confidence intervals shown.

Figure A34: Prob. homicide by mean scale



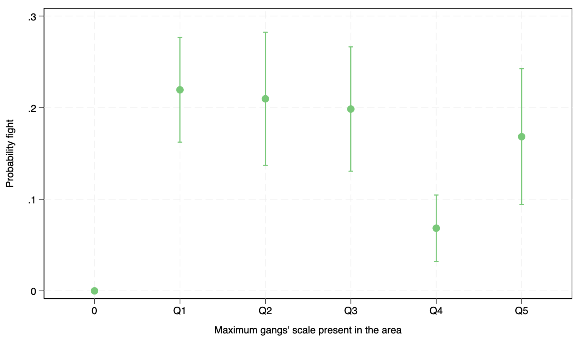
Notes: The figure shows the average probability of gang homicides for location-years with no gangs and for five quintiles of the average scale of gangs present in the area. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. The scale of a gang, in a given year, is the number of locations in which it is present. Each observation is a location-year. Locations are districts within Naples and cities in the province. 95% confidence intervals shown.

Figure A35: Prob. gang homicide by mean scale



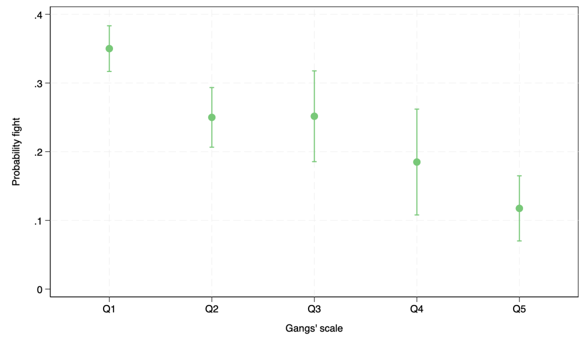
Notes: The figure shows the average probability of non-gang homicides for location-years with no gangs and for five quintiles of the average scale of gangs present in the area. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. The scale of a gang, in a given year, is the number of locations in which it is present. Each observation is a location-year. Locations are districts within Naples and cities in the province. 95% confidence intervals shown.

Figure A36: Prob. non-gang homicide by mean scale



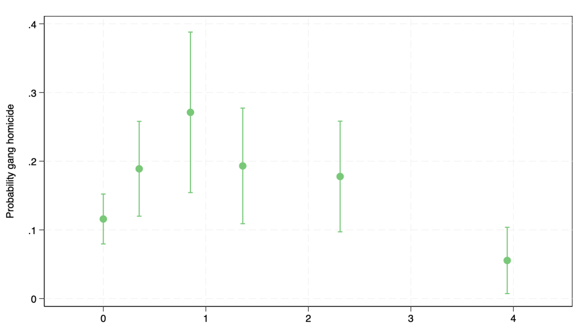
Notes: The figure shows the average probability of fights for location-years with no gangs and for five quintiles of the maximum scale of gangs present in the area. The scale of a gang, in a given year, is the number of locations in which it is present. Each observation is a location-year. Locations are districts within Naples and cities in the province. 95% confidence intervals shown.

Figure A37: Prob.fight by maximum scale



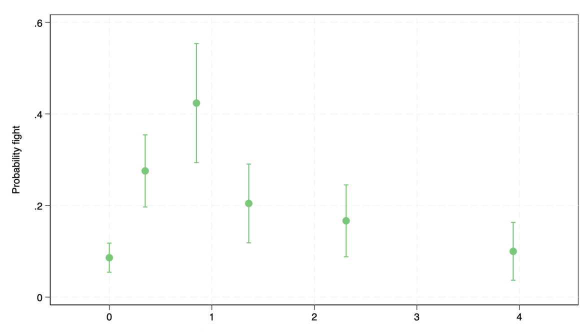
Notes:The figure shows the average probability of fights for gang-years with different scale. The scale of a gang, in a given year, is the number of locations in which it is present. Each observation is a gang-year. Locations are districts within Naples and cities in the province. 95% confidence intervals shown.

Figure A38: Prob. fight by scale



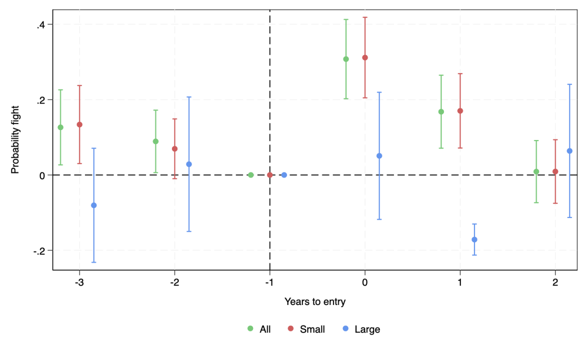
Notes: The figure shows the average probability of gang homicides for location-years with no gangs and for five quintiles of the standard deviation of active gangs. Each observation is a location-year. A location is a district within Naples, and a city in the province. 95% confidence intervals shown.

Figure A39: Prob. gang homicide by sd scale



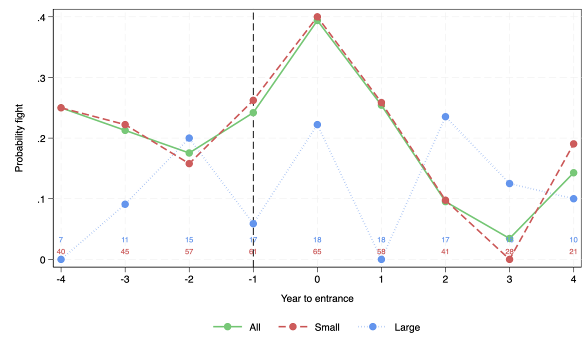
Notes: The figure shows the average fight probability for location-years with no gangs and for five quintiles of the standard deviation of active gangs. Each observation is a location-year. A location is a district within Naples, and a city in the province. 95% confidence intervals shown.

Figure A40: Prob. fight by sd scale



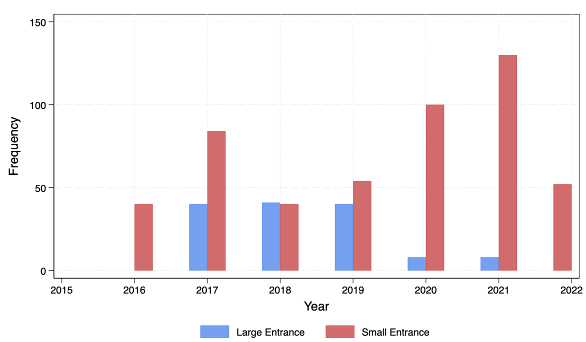
Notes: This figure presents the event-study coefficients separately for the last entry of a new gang in the location, the last entry of a small gang, and the last entry of a large gang, on the probability of fights in the location. Each observation is a location-year. Sample restricted to location-years with at least one active gang. A gang is defined to be small if it has a scale of 3 or lower (90th scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to entry in an area at time t if is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 90% confidence interval shown. Standard errors clustered at the location level. A location is a district within Naples, and a city in the province.

Figure A41: Gang entry and probability of fight



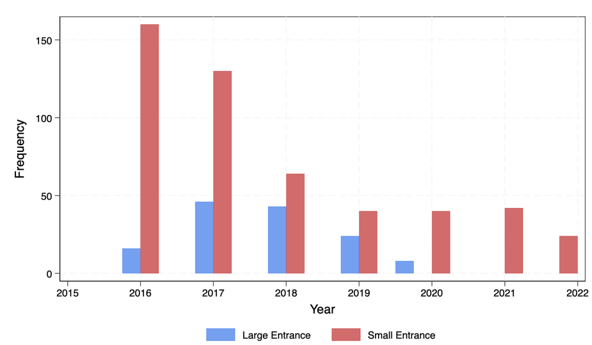
Notes:The figure shows the average probability of fights for location-years relative to the last entry of a new gang in the area, separately for general entry, small gang entry, and large gang entry. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Locations are districts within Naples and cities in the province.

Figure A42: Event study - probability fights, raw data



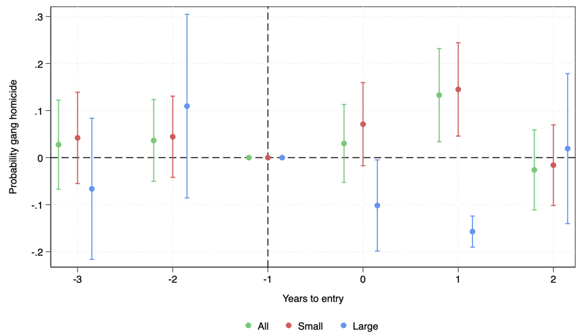
Notes: The figure shows the frequency of last entry in an area, separately for large and small gangs. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Locations are districts within Naples and cities in the province.

Figure A43: Event study - variation entrance



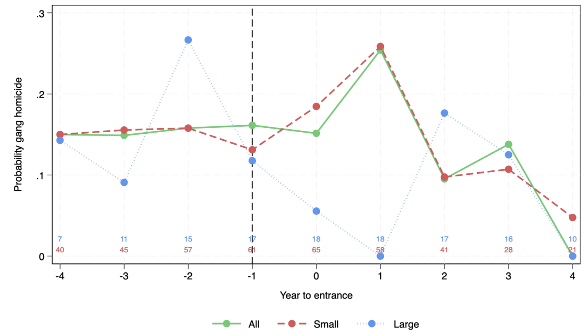
Notes: The figure shows the frequency of first entry in an area, separately for large and small gangs. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Locations are districts within Naples and cities in the province.

Figure A44: Event study - variation entrance, first entry



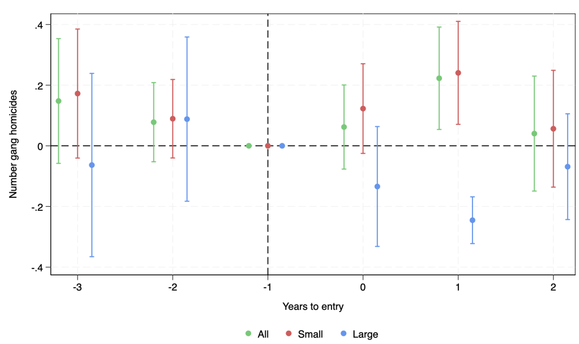
Notes: The figure shows event-study coefficients of separate regressions for the last entry of a new gang in the area, the last entry of a small gang, and the last entry of a large gang, on the probability of gang homicide. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 90% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A45: Event study entry - probability gang homicide



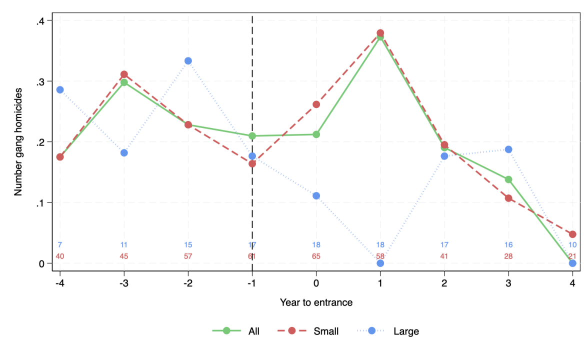
Notes: The figure shows the average probability of gang homicides for location-years relative to the last entry of a new gang in the area separately for general entry, small gang entry, and large gang entry. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Locations are districts within Naples and cities in the province.

Figure A46: Event study entry - probability gang homicide, raw data



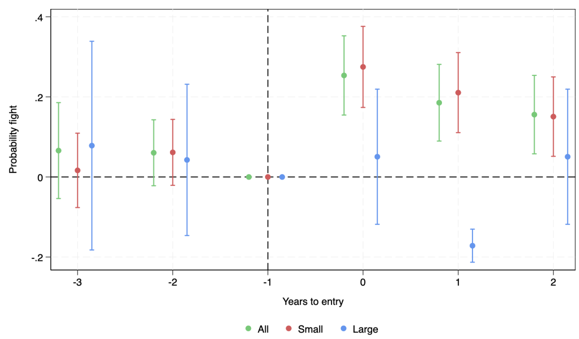
Notes: The figure shows event-study coefficients of separate regressions for the last entry of a new gang in the area, the last entry of a small gang, and the last entry of a large gang, on the number of gang homicides. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 90% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A47: Event study entry - number gang homicides



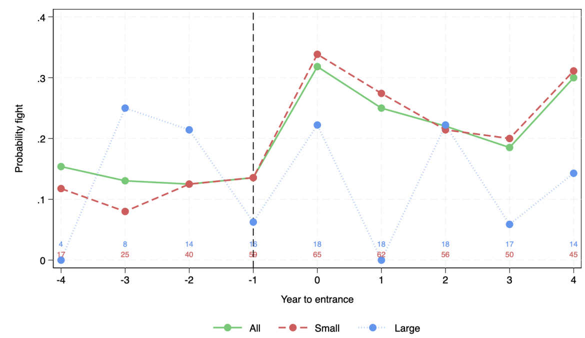
Notes: The figure shows the average number of gang homicides for location-years relative to the last entry of a new gang in the area separately for general entry, small gang entry, and large gang entry. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Locations are districts within Naples and cities in the province.

Figure A48: Event study entry - number gang homicides, raw data



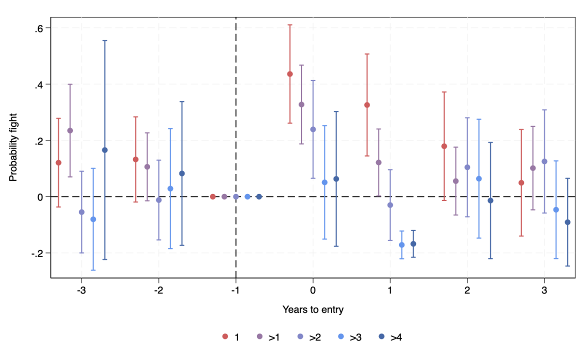
Notes: The figure shows event-study coefficients of separate regressions for the first entry of a new gang in the area, the first entry of a small gang, and the first entry of a large gang, on the probability of fights. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 90% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A49: Event study entry - probability fights, first entry



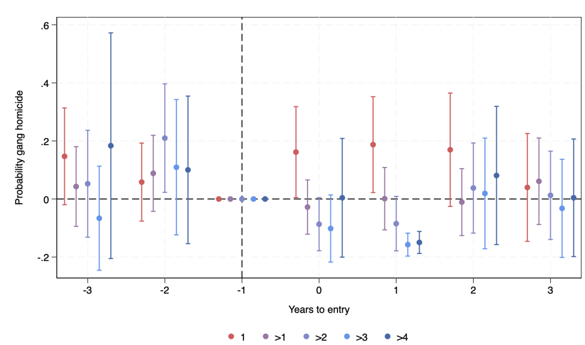
Notes: The figure shows the average probability of fights for location-years relative to the first entry of a new gang in the area of separate regressions for general entry, small gang entry, and large gang entry. Each observation is a location-year. Sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Locations are districts within Naples and cities in the province.

Figure A50: Event study entry - probability fights, first entry raw data



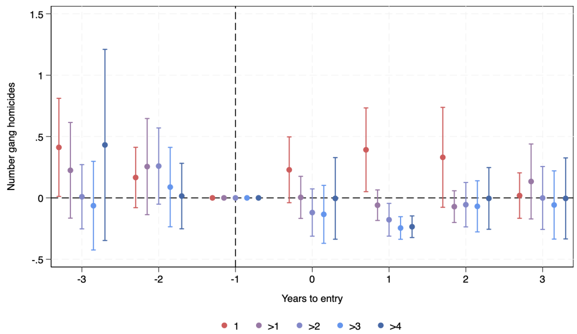
Notes: The figure shows event-study coefficients of separate regressions for the last entry of gangs of different scale on the probability of fights in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A51: Event study entry - fight, robustness definition large



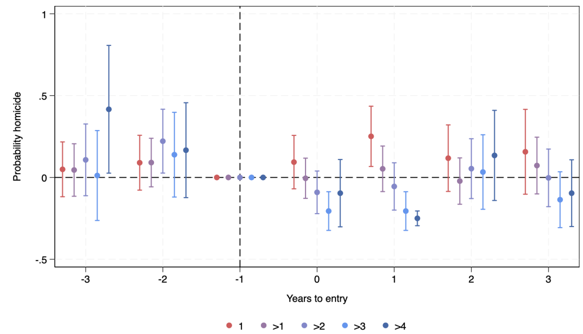
Notes: The figure shows event-study coefficients of separate regressions for the last entry of gangs of different scale on the probability of gang homicide in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A52: Event study entry - probability gang homicide, robustness definition large



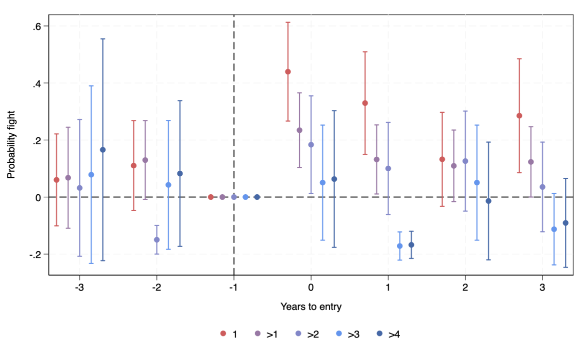
Notes: The figure shows event-study coefficients of separate regressions for the last entry of gangs of different scale on the number of gang homicides in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A53: Event study entry - number gang homicide, robustness definition large



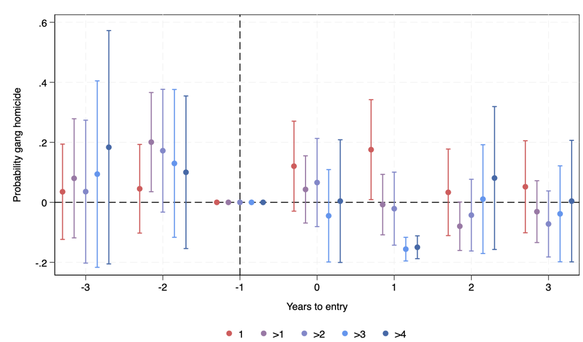
Notes: The figure shows event-study coefficients of separate regressions for the last entry of gangs of different scale on the probability of homicide in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A54: Event study entry - probability homicide, robustness definition large



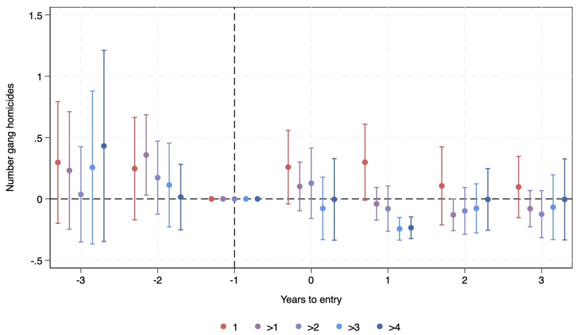
Notes: The figure shows event-study coefficients of separate regressions for the first entry of gangs of different scale on the probability of fights in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A55: Event study entry - fight, robustness definition large, first entry



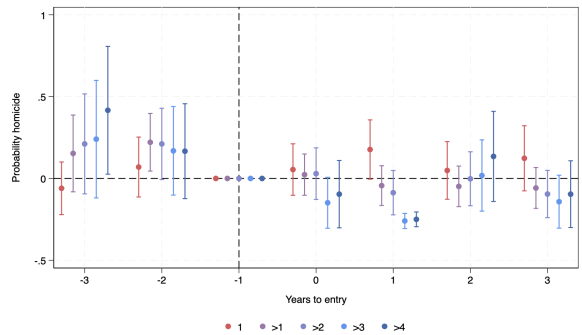
Notes: The figure shows event-study coefficients of separate regressions for the first entry of gangs of different scale on the probability of gang homicide in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A56: Event study entry - probability gang homicide, robustness definition large



Notes: The figure shows event-study coefficients of separate regressions for the first entry of gangs of different scale on the number of gang homicides in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A57: Event study entry - number gang homicide, robustness definition large



Notes: The figure shows event-study coefficients of separate regressions for the first entry of gangs of different scale on the probability of homicide in the area. Each observation is a location-year. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A58: Event study entry - probability homicide, robustness definition large

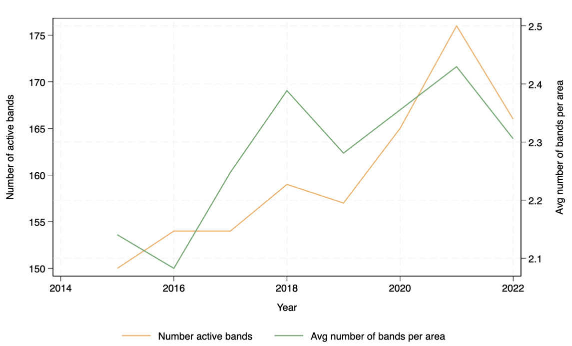


Figure A59: Number of active gangs

Notes: The figure shows number of active gangs and the average number of gangs per location for each year of the sample.

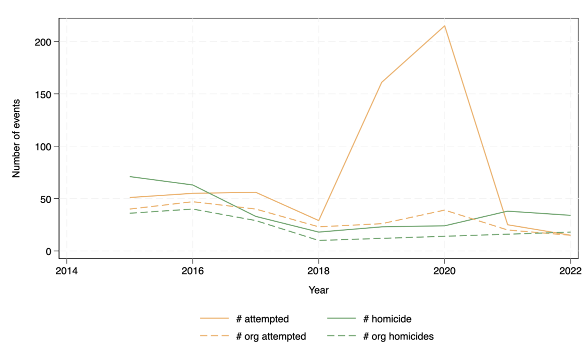
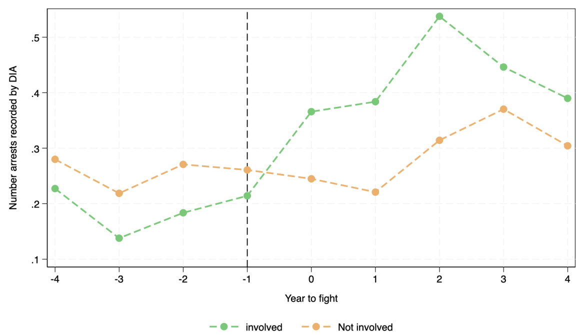


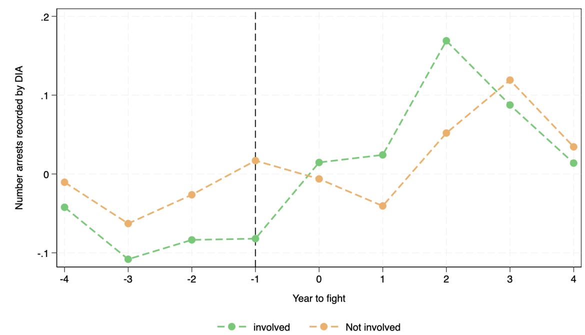
Figure A60: Time trend violence

Notes: The figure shows number of attempted and total homicides, gang related or not, in each year of the sample.



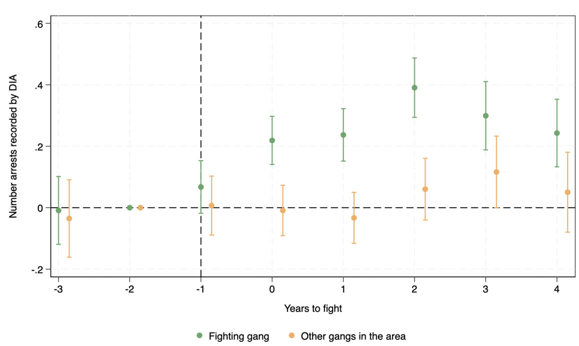
Notes: The figure shows the average number of arrests recorded in DIA reports relative to the first fight. Each observation is a location-gang-year. The average number of DIA-reported arrests for each gang around its first fight is reported in green. In yellow it is reported the average number of DIA-reported arrests in an area where the gang was present but not directly involved. Locations are districts within Naples and cities in the province.

Figure A61: Event study spillover - data



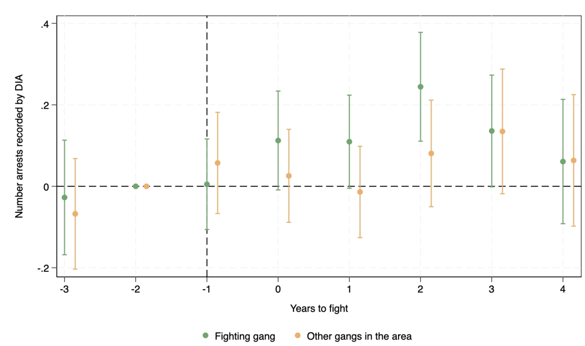
Notes: The figure shows the average number of arrests recorded in DIA reports relative to the first fight. Each observation is a location-gang-year. The average number of DIA-reported arrests for each gang around its first fight is reported in green. In yellow it is reported the average number of DIA-reported arrests in an area where the gang was present but not directly involved. Locations are districts within Naples and cities in the province. Gang fixed effects are included.

Figure A62: Event study spillover - data, fixed effects



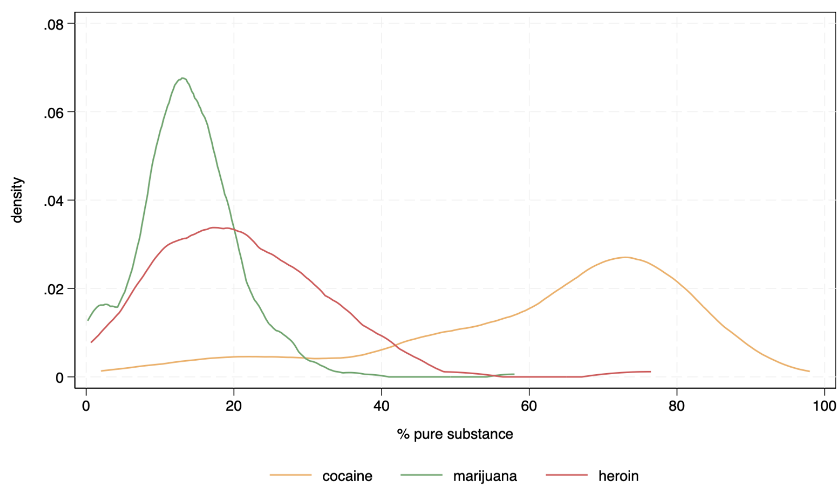
Notes: The figure shows event-study coefficients of separate regressions for the first fight of a gang (in green) and the first fight in a location where the gang was present but not directly involved in the fight (in yellow). Each observation is a location-gang-year. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province.

Figure A63: Event study spillover



Notes: The figure shows event-study coefficients of separate regressions for the first fight of a gang (in green) and the first fight in a location where the gang was present but not directly involved in the fight (in yellow). Each observation is a location-gang-year. EVENTDD estimates Clarke & Tapia Schythe, 2023. 95% confidence intervals shown. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province. Gang fixed effects are included.

Figure A64: Event study spillover, fixed effects



Notes: The figure shows the distribution of percentage of pure substance in the seizure data, separately by drug type.

Figure A65: Distribution drug quality

Table A21: Gang entry

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: probability fight						
Entry small	0.346***		0.221***		0.198***	
	(0.0454)		(0.0450)		(0.0459)	
Entry large		0.00137		0.0757		0.0623
		(0.0815)		(0.0737)		(0.0759)
Observations	756	756	753	753	753	753
R-squared	0.119	0.0000645	0.379	0.337	0.389	0.356
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.109	0.165	0.109	0.165	0.109	0.165
Pvalue equality	0.000		0.067		0.129	
Panel B: probability gang homicide						
Entry small	0.0837**		0.00102		0.00634	
	(0.0398)		(0.0384)		(0.0382)	
Entry large		-0.0683		-0.0111		0.0163
		(0.0558)		(0.0510)		(0.0534)
Observations	756	756	753	753	753	753
R-squared	0.00750	0.00113	0.245	0.245	0.267	0.267
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.135	0.152	0.135	0.152	0.135	0.152
Pvalue equality	0.006		0.837		0.882	
Panel C: number gang homicide						
Entry small	0.170**		0.0102		0.0147	
	(0.0789)		(0.0640)		(0.0588)	
Entry large		-0.110		-0.0111		0.0313
		(0.0935)		(0.0793)		(0.0846)
Observations	756	756	753	753	753	753
R-squared	0.00832	0.000784	0.324	0.324	0.351	0.351
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.204	0.235	0.204	0.235	0.204	0.235
Pvalue equality	0.019		0.829		0.889	
Panel D: average gang scale						
Entry small	-0.781***		-0.251***		-0.244***	
	(0.167)		(0.0641)		(0.0682)	
Entry large		0.568***		0.519***		0.465***
		(0.216)		(0.124)		(0.118)
Observations	756	756	753	753	753	753
R-squared	0.0338	0.00403	0.845	0.845	0.851	0.851
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	3.180	3.035	3.180	3.035	3.180	3.035
Pvalue equality	0.000		0.000		0.000	
Panel E: number gangs						
Entry small	1.778***		0.779***		0.806***	
	(0.251)		(0.0903)		(0.0890)	
Entry large		1.031***		0.336		0.306
		(0.336)		(0.269)		(0.257)
Observations	756	756	753	753	753	753
R-squared	0.0986	0.00749	0.913	0.899	0.916	0.902
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	2.628	2.885	2.628	2.885	2.628	2.885
Pvalue equality	0.007		0.020		0.013	

	(1)	(2)	(3)	(4)	(5)	(6)
Panel F: probability fight, including period after entry						
Entry small	0.305***		0.194***		0.185***	
	(0.0452)		(0.0473)		(0.0466)	
Entry large		-0.0813*		-0.0360		-0.0466
		(0.0455)		(0.0414)		(0.0430)
Observations	756	756	753	753	753	753
R-squared	0.138	0.000778	0.380	0.336	0.393	0.355
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.081	0.170	0.081	0.170	0.081	0.170
Pvalue equality	0.000		0.000		0.000	
Panel G: probability gang homicide, including period after entry						
Entry small	0.107***		0.0318		0.0690*	
	(0.0390)		(0.0392)		(0.0391)	
Entry large		-0.112***		-0.0810**		-0.0441
		(0.0328)		(0.0406)		(0.0427)
Observations	756	756	753	753	753	753
R-squared	0.0179	0.00549	0.246	0.247	0.271	0.267
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.120	0.156	0.120	0.156	0.120	0.156
Pvalue equality	0.000		0.055		0.055	
Panel H: number gang homicide, including period after entry						
Entry small	0.174**		0.00419		0.0681	
	(0.0664)		(0.0812)		(0.0725)	
Entry large		-0.175***		-0.117**		-0.0345
		(0.0595)		(0.0575)		(0.0607)
Observations	756	756	753	753	753	753
R-squared	0.0126	0.00363	0.324	0.325	0.353	0.351
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.184	0.241	0.184	0.241	0.184	0.241
Pvalue equality	0.000		0.269		0.325	
Panel I: average gang scale, including period after entry						
Entry small	-0.797***		-0.229***		-0.263***	
	(0.191)		(0.0783)		(0.0844)	
Entry large		0.438*		0.426***		0.336***
		(0.234)		(0.0955)		(0.0996)
Observations	756	756	753	753	753	753
R-squared	0.0515	0.00436	0.845	0.845	0.852	0.850
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	3.271	3.027	3.271	3.027	3.271	3.027
Pvalue equality	0.000		0.000		0.000	
Panel L: number gangs, including period after entry						
Entry small	1.823***		0.722***		0.719***	
	(0.275)		(0.0813)		(0.0804)	
Entry large		1.150***		0.573***		0.543***
		(0.370)		(0.186)		(0.188)
Observations	756	756	753	753	753	753
R-squared	0.151	0.0170	0.914	0.901	0.915	0.904
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	2.418	2.849	2.418	2.849	2.418	2.849
Pvalue equality	0.008		0.055		0.067	

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Entry small is a dummy variable equal to one when a small gang enters the location. Entry large is a dummy variable equal to one when a large gang enters the location. In Panels F-L both dummy variables are equal to one also the period after the entry. An homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide.

Table A22: Gang entry - sensitivity, min scale

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: probability fight						
Entry small min	0.335***		0.212***		0.190***	
	(0.0451)		(0.0441)		(0.0449)	
Entry large min		-0.0158		0.0534		0.0477
		(0.0840)		(0.0652)		(0.0734)
Observations	756	756	753	753	753	753
R-squared	0.114	0.000	0.377	0.337	0.387	0.355
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.109	0.165	0.109	0.165	0.109	0.165
Pvalue equality	0.000		0.049		0.122	
Panel B: probability gang homicide						
Entry small min	0.0778**		-0.00100		0.00446	
	(0.0387)		(0.0371)		(0.0373)	
Entry large min		-0.0508		0.00534		0.0398
		(0.0698)		(0.0687)		(0.0680)
Observations	756	756	753	753	753	753
R-squared	0.00661	0.000	0.245	0.245	0.267	0.267
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.136	0.149	0.136	0.149	0.136	0.149
Pvalue equality	0.041		0.930		0.645	
Panel C: number gang homicide						
Entry small min	0.160**		0.00802		0.0137	
	(0.0768)		(0.0624)		(0.0574)	
Entry large min		-0.0837		0.00534		0.0548
		(0.113)		(0.100)		(0.101)
Observations	756	756	753	753	753	753
R-squared	0.00753	0.000	0.324	0.324	0.351	0.351
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.204	0.232	0.204	0.232	0.204	0.232
Pvalue equality	0.054		0.976		0.747	

	(1)	(2)	(3)	(4)	(5)	(6)
Panel D: probability fight, including period after entry						
Entry small min	0.297***		0.187***		0.178***	
	(0.0448)		(0.0465)		(0.0455)	
Entry large min		-0.0910**		-0.0577*		-0.0635*
		(0.0458)		(0.0301)		(0.0342)
Observations	756	756	753	753	753	753
R-squared	0.133	0.000	0.378	0.336	0.391	0.355
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.081	0.170	0.081	0.170	0.081	0.170
Pvalue equality	0.000		0.000		0.000	
Panel E: probability gang homicide, including period after entry						
Entry small min	0.102***		0.0297		0.0680*	
	(0.0384)		(0.0384)		(0.0384)	
Entry large min		-0.102***		-0.0793		-0.0400
		(0.0386)		(0.0520)		(0.0533)
Observations	756	756	753	753	753	753
R-squared	0.0164	0.004	0.246	0.247	0.271	0.267
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.121	0.154	0.121	0.154	0.121	0.154
Pvalue equality	0.000		0.100		0.097	
Panel F: number gang homicide, including period after entry						
Entry small min	0.165**		0.00247		0.0693	
	(0.0653)		(0.0799)		(0.0713)	
Entry large min		-0.161**		-0.123*		-0.0373
		(0.0669)		(0.0711)		(0.0706)
Observations	756	756	753	753	753	753
R-squared	0.0116	0.003	0.324	0.325	0.353	0.351
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.185	0.239	0.185	0.239	0.185	0.239
Pvalue equality	0.000		0.285		0.324	

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. A gang is defined as small if it has a scale of 3 or lower (90th percentile of scale distribution). The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Entry small min is a dummy variable equal to one when a gang with minimum scale (over all time periods) lower or equal than 3 enters the location. Entry large min is a dummy variable equal to one when a gang with minimum scale (over all time periods) higher than 3 enters the location. In Panels D-F both dummy variables are equal to one also the period after the entry. An homicide is coded as gang homicide if there was any gang mentioning on the internet related to the homicide.

Table A23: Gang entry - sensitivity, relative entry

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: probability fight						
Entry decreasing	0.297***		0.210***		0.185***	
	(0.0521)		(0.0514)		(0.0522)	
Entry increasing		0.297***		0.164**		0.148*
		(0.0871)		(0.0810)		(0.0802)
Observations	756	756	753	753	753	753
R-squared	0.114	0.000	0.377	0.337	0.387	0.355
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.123	0.148	0.123	0.148	0.123	0.148
Pvalue equality	0.787		0.605		0.711	
Panel B: probability fight, including period after entry						
Entry decreasing	0.285***		0.206***		0.195***	
	(0.0476)		(0.0459)		(0.0457)	
Entry increasing		0.211***		0.0713		0.0617
		(0.0672)		(0.0685)		(0.0654)
Observations	756	756	753	753	753	753
R-squared	0.00661	0.000	0.245	0.245	0.267	0.267
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.095	0.142	0.095	0.142	0.095	0.142
Pvalue equality	0.084		0.041		0.065	

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: probability fight						
Entry decreasing min	0.313***		0.222***		0.198***	
	(0.0496)		(0.0478)		(0.0480)	
Entry increasing min		0.224**		0.108		0.0998
		(0.0908)		(0.0848)		(0.0864)
Observations	756	756	753	753	753	753
R-squared	0.133	0.000	0.378	0.336	0.391	0.355
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.119	0.154	0.119	0.154	0.119	0.154
Pvalue equality	0.261		0.273		0.386	
Panel D: probability fight, including period after entry						
Entry decreasing min	0.290***		0.212***		0.202***	
	(0.0476)		(0.0459)		(0.0452)	
Entry increasing min		0.146**		0.00956		0.00201
		(0.0719)		(0.0746)		(0.0725)
Observations	756	756	753	753	753	753
R-squared	0.0164	0.004	0.246	0.247	0.271	0.267
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep.	0.092	0.152	0.092	0.152	0.092	0.152
Pvalue equality	0.012		0.007		0.013	

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Entry decreasing is a dummy variable equal to one when a gang with scale lower than the average of the location enters it. Entry increasing is a dummy variable equal to one when a gang with scale higher than the average of the location enters it. Entry decreasing min is a dummy variable equal to one when a gang with minimum scale (over all time periods) enters in a location with lower or equal average scale. Entry increasing min is a dummy variable equal to one when a gang with minimum scale (over all time periods) enters in a location with higher average scale.

Table A24: Gang entry - sensitivity, definition

Gang scale	Sum of locations		Sum of population		Sum of consumer's share		Sum of seized share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Entry small	0.208***		0.181***		0.190***		0.164***	
	(0.0496)		(0.0441)		(0.0442)		(0.0427)	
Entry large		0.0623		0.117		0.0861		0.270**
		(0.0759)		(0.0983)		(0.125)		(0.105)
Observations	753	753	753	753	753	753	753	753
Location FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Entry small is a dummy variable equal to one when a small gang enters the location. Entry large is a dummy variable equal to one when a large gang enters the location. A gang is defined to be large (and small) if its scale is higher (or lower) than the 90th percentile of the gang scale distribution. In columns (1)-(2) gang scale is measured as the number of locations in which the gang is present. In (3)-(4) scale is measured as the (standardized) total number of people living in the locations where the gang is present. In (5)-(6) scale is measured as the number of locations where the gang is present, weighted by the average share of estimated consumers across drugs. In (7)-(8) is measured as the number of locations where the gang is present, weighted by the share of seizures from that location in the sample.

Table A25: Gang entry - robustness large definition

Dep. Variable: Probability	(1) Fight	(2) Fight	(3) Fight	(4) Fight	(5) Fight	(6) Fight
Entry scale >0	0.198***					
	(0.0691)					
Entry scale >1		0.193***				
		(0.0567)				
Entry scale >2			0.194***			
			(0.0722)			
Entry scale >3 (benchmark)				0.0623		
				(0.0759)		
Entry scale >4					0.0851	
					(0.106)	
Entry scale >5						0.0861
						(0.125)
Observations	753	753	753	753	753	753
R-squared	0.368	0.384	0.363	0.356	0.356	0.356
Location FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep | No entry	0.139	0.139	0.139	0.139	0.139	0.139

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. The scale of a gang, in a given year, is the number of locations in which it is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Entry scale ¿ x is a dummy variable equal to one when a gang with scale higher than x enters the location.

Table A26: Gang entry - entry in another location

Dep. Variable: Probability	(1) Fight	(2) Fight	(3) Fight	(4) Fight	(5) Fight	(6) Fight
Panel A: OLS						
Average scale	-0.0579***	-0.0341***	-0.0599***	-0.0401***	-0.0549***	-0.0328**
	(0.0110)	(0.00942)	(0.0157)	(0.0138)	(0.0165)	(0.0136)
Number gangs		0.0533***		0.0937***		0.0955***
		(0.00883)		(0.0219)		(0.0218)
Observations	756	756	756	753	753	753
R-squared	0.0597	0.139	0.338	0.365	0.355	0.382
Location FE	No	No	No	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Mean Dep | Average scale = 1	0.226	0.226	0.226	0.226	0.226	0.226
Panel B: Reduced form						
Entry in another location	0.0637	-0.0117	-0.0174	-0.0337	-0.0314	-0.0409
	(0.0469)	(0.0400)	(0.0369)	(0.0371)	(0.0407)	(0.0394)
Number gangs		0.0624***		0.102***		0.103***
		(0.00915)		(0.0230)		(0.0223)
Observations	756	756	753	753	753	753
R-squared	0.00438	0.121	0.328	0.361	0.348	0.380
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep | No entry	0.155	0.155	0.155	0.155	0.155	0.155
Panel C: First stage						
Entry in another location	0.344	0.682***	0.507***	0.539***	0.489***	0.508***
	(0.2187)	(0.1950)	(0.1020)	(0.1044)	(0.0948)	(0.0995)
Number gangs		-0.280***		-0.202***		-0.208***
		(0.0540)		(0.0802)		(0.0832)
Observations	756	756	753	753	753	753
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Panel D: Second stage						
Average scale	0.185	-0.0171	-0.0344	-0.0625	-0.0643	-0.0805
	(0.203)	(0.0582)	(0.0741)	(0.0720)	(0.0865)	(0.0822)
Number gangs		0.0576***		0.0896***		0.0860***
		(0.0173)		(0.0218)		(0.0233)
Observations	756	756	753	753	753	753
R-squared	-0.993	0.135	0.0122	0.0533	0.0120	0.0449
Location FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	Yes	Yes
Mean Dep | Average scale == 1	0.226	0.226	0.226	0.226	0.226	0.226
KP-stat	2.468	12.22	24.67	26.65	26.52	26.04

Notes: HDFE linear regression. Each observation is a location-year. Standard errors clustered at the location level. ***,**,* = indicate significance at the 1, 5, and 10% level, respectively. Locations are districts within Naples and cities in the province. Sample is restricted to observations with at least one active gang. Average scale is the average scale of gangs in the location, measured as the number of locations in which the gang is present. A gang is defined to enter an area at time t if it is present in the area at t, but not in $t-1$. Entry in another location is a dummy variable equal to one for observations in which an active gang has entered another location, including the period after entry. This is missing for observations in which the only entry was in the location itself. Panel A shows the simple OLS regression of average scale and number of active gangs on the probability of fights. Panel B OLS regression of entry in another location and number gangs on the probability of fights. Panel C OLS regressions of entry in another location and number of gangs on the average gang scale in the area. Panel D 2SLS regressions of average scale, instrumented with entry in another location, and the number of gangs on the probability of fights.