Collegio Carlo Alberto

Understanding Organized Crime Networks: Evidence Based on Federal Bureau of Narcotics Secret Files on American Mafia

Giovanni Mastrobuoni
Eleonora Patacchini

No. 152
September 2010

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

Understanding Organized Crime Networks: Evidence Based on Federal Bureau of Narcotics Secret Files on American Mafia *

Giovanni Mastrobuoni[†] and Eleonora Patacchini[‡]

September 2010[§]

*We would like to thank Theo Diasakos, Jim Heckman, Matthew Jackson, Claudio Lucifora, Franco Peracchi, Rocco Sciarrone, Serena Uccello, Aleksey Tetenov and seminar participants at the Workshop on the "Economics of Crime and Organized Crime" in Palermo, the one on "Institutions, Individual Behavior and Economic Outcomes" in Alghero, and the one in Petralia for their useful comments. Martino Bernardi, Isabella David, Filippo Maggi, and Dominic Smith have provided excellent research assistance. This research was supported by a Collegio Carlo Alberto grant. Giovanni Mastrobuoni thanks the Italian Academy at Columbia University for their hospitality.

[†]Corresponding author. Associate Research Scholar, Italian Acedemy, Columbia University, Collegio Carlo Alberto and CeRP, Via Real Collegio 30, Moncalieri, Italy, giovanni.mastrobuoni@carloalberto.org.

[‡]Department of Economic and Social Analysis, Università la Sapienza, EIEF, IZA, and CEPR, P.le A. Moro 5 - 00185 Roma, Italy, eleonora.patacchini@uniroma1.it

Abstract

Using unique data on criminal profiles of 800 US Mafia members active in the 50s and 60s and on their connections within the *Cosa Nostra* network we analyze how the geometry of criminal ties between mobsters depends on family ties, community roots and ties, legal and illegal activities. We contrast our evidence with historical and sociological views about the functioning of the Mafia. Much of our findings are remarkably in line with these views, with interesting qualifications. We interpret some of our results in light of a model of optimal vertical and horizontal connections where more connections mean more profits but also a higher risk of defection. We find that variables that lower the risk of defection, among others, kinship, violence, and mafia *culture* increase the number of connections. Moreover, there is evidence of strategic endogamy: female children are as valuable as male ones, and being married to a "connected" wife is a strong predictor of leadership within the Mafia ranks. A very parsimonious regression model explains one third of the variability in the criminal ranking of the "men of honor," suggesting that these variables could be used to detect criminal leaders. An additional prediction of our simple model is a right-skewed distribution of the number of connections, which is remarkably in line with the evidence of an extremely hierarchical organization.

Keywords: Mafia, Networks, Intermarriage, Assortative Matching, Crime.
JEL classification codes: A14, C21, D23, D85, K42, Z13

1 Introduction

Despite more than 40 years of law enforcement's success in the fight against organized crime, according to the FBI[1], in 2005 there were 651 pending investigations related to the Italian Mafia in the U.S.; almost 1,500 mobsters were arrested, and 824 were convicted; of the roughly 1,000 "made" members of Italian organized crime groups estimated to be active in the U.S., 200 were in jail. In addition, the Italian Mafia doesn't anymore hold the monopoly of racketeering. With the end of the Cold War and the advent of globalization "transnational" organized crime organization are on the rise, mainly the Russian Mafia, the African enterprises, the Chinese tongs, South American drug cartels, the Japanese Yakuza, and the, so called, Balkan Organized Crime groups, and their proceeds, by the most conservative estimates, comprise around 5 percent of the world's GDP (Schneider and Enste, 2000, Wagley, 2006).

Notwithstanding such numbers, the illicit nature of organized crime activities has precluded quantitative analysis and the literature has overwhelmingly been anecdotal or theoretical (Reuter, 1994). Indeed, this is the first study that uses extensive individual level data to describe the hierarchy of such organizations.[2] We use criminal records for more than 800 criminals–Figure 1 shows record number one—-based on an exact facsimile of the Federal Bureau of Narcotics (FBN) secret files on American Mafia members that were active and alive in 1960 (MAF, 2007) to analyze criminal connections and hierarchies, ranking mobsters based on the number and the quality of their connections.[3] A possible policy purpose is to learn about organizational rules of the Mafia and to identify the "key players."

[1]The source is www.fbi.gov.

[2]A rare exception is Levitt and Venkatesh (2000) who use detailed financial activities of a drug-selling street gang to analyze gang behavior.

[3]At the end of the 50s the FBN, that later merged with the Bureau of Drug Abuse Control to form the Bureau of Narcotics and Dangerous Drugs, was the main authority in the fight against the Mafia (Critchley, 2009). In New York the FBI had just four agents, mainly working in office, assigned to the area, while in the same office more than 400 agents were fighting domestic communists (Maas, 1968).

We use several measures on how central members are within the network. A mobster might exert power because he is largely connected, or because he has few connections but these are to high caliber figures like Lucky Luciano, Frank Costello, or Joe Bonanno. Mobsters might also be central because they represent bridges that connect different clusters of a network. The geometry of Mafia connections is crucial for understanding the activity of the *Cosa Nostra*, "Our Thing," as these connections are the building block of the entire Mafia, and more generally of organized crime, even today. Valachi's 1962 testimony and documents found during the 2007 arrest of Salvatore Lo Piccolo, a Sicilian Mafia boss, show that the first rule in the Mafia decalogue stays unchallenged: "No one can present himself directly to another of our friends. There must be a third person to do it" (Maas, 1968). Connections are thus necessary for a criminal career within the Mafia. Moreover, leadership positions cannot be simply inherited; soldiers elect their boss using secret ballots (Falcone and Padovani, 1991, pg. 101).

Francisco Costiglia, alias Frank Costello, a Mafia bosses who in our data is connected to 34 mobsters, would say "he is connected" to describe someone's affiliation to the Mafia (Wolf and DiMona, 1974). In 1970 the Organized Crime Control Act defined organized crime as "The unlawful activities of ... a highly organized, disciplined association...." The purpose of our research is to describe these associations and their organizational structure. The structure has its roots in a world characterized by the absence of legally enforceable contracts. These criminals need to trust each other, and the purpose of this study is to understand how and where this trust emerges. To the best of our knowledge no one has used network analysis tools on such a detailed set of information on individual Mafia members to study the emergence of networks: ranging from their business to their family structure.

Guided by a simple model of connections we try to shed light on several predictions about how these networks and the related hierarchies emerge. We evaluate the relative

importance of legal and illegal businesses, family ties, and community ties in shaping these networks, contrasting the economic with the social view of *Cosa Nostra*.

Criminals might be more likely to be associated with criminals who establish similar illegal businesses if they try to build cartels. Or they might try to diversify the risk of detection keeping a lower profile and thus associate themselves with criminals who operate different kinds of businesses. Carefully chosen marriages might help to establish robust criminal ties. Children might thus be important too, both because of this strategic endogamy, and because male descendants represent trusted potential associates.[4] [5] Based on a participant observation study for a New York based Italian American crime family Ianni and Reuss-Ianni (1972) discuss the importance of family lineage, intermarriage, and kinship.

> As in the great *Mafia* families of Sicily, the "descendants" intermarry continuously, for the clan defines who may or may not marry whom. As in clans everywhere, the relationship between the intermarrying pair is defined as strengthening the social structure. (Ianni and Reuss-Ianni, 1972, pg. 192)

These marriages would thus have the same function as alliances among European royalty: providing new protectors.

Another way children or, more generally, relatives might increase the trust toward a member is because they represent potential targets for retaliatory action. Trust, or better, blind obedience, and the vow of silence, called *omertà*, are indeed essential for the Mafia clans' survival. Mafia clans, called "Families" (as in Bonanno (1983) we use uppercase to distinguish them from the nuclear family) represent societies where social capital produces public "bads" (Portes, 1998). A Family protects its members and guarantees

[4] Information on the number of male and female children, and on the maiden name of the wife is going to allow us to verify the strategic endogamy theory and to evaluate the importance of descendants.

[5] Another instrument for building bonds is the "comparatico," a spiritual parentage a la "The Godfather." Unfortunately the data do not contain information about these kind of links (Ianni and Reuss-Ianni, 1972).

their monopoly power in exchange of part of their revenues. Large clans will therefore be more powerful but also more exposed. A simple model of sequential "dictatorial" network formation by a Mafia boss formalizes this tradeoff and guides our empirical strategy.

Some of the economic insights are present in early theoretical analysis of criminal behavior. But most studies have focused on a market structure view of organized crime, where the Mafia generates monopoly power in legal (for a fee) and illegal markets. Among others, such view is present in the collection of papers in Fiorentini and Peltzman (1997), and in Reuter (1983), Abadinsky (1990), Gambetta (1996), and Kumar and Skaperdas (2009). Only two theoretical papers have focused on the internal organization of organized crime groups. Garoupa (2007) looks at the optimal size of these organizations, while Baccara and Bar-Isaac (2008) looks at the optimal internal structure (cells versus hierarchies). We borrow some of their insights, mainly that larger organizations are more profitable but also more vulnerable.

Our work is also related to the growing literature on trust, family values, and family businesses. Guiso et al. (2006) present an introduction to the importance of culture, defined as "customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation," on economic behavior. We will argue that for criminal behavior the same applies. Bertrand and Schoar (2006) present a macro-type analysis about the importance of family values for economic growth, and conclude writing that more research is need to understand how family values shape the organization of businesses and their efficiency.

Sparrow (1991) proposes the use of network analysis to study criminal networks, but only Morselli (2003) followed Sparrow's proposal, and only based on a single New York based family, the Gambino one. Recent papers on social networks show that the individual position within a network is indeed crucial in explaining the individual level of activity (Ballester et al., 2006). If individual decisions are somehow connected to the

structure of social contacts each individual chooses (or is trapped in), understanding the formation of network structure is crucial for anti-crime policies. Given the complexity of social relationship, evidence documenting patterns of association between agents is a first priority as it can also inform the theoretical literature looking at network formation processes. Several models of network formation have recently been proposed. Most rely on some forms of pairwise regressions, but are based on strong assumptions (see, for example, Bramoullé and Fortin, 2010). We focus more on the elements that are related to the centrality of mobsters than on those that determine single connections.

In this paper, we contribute to this strand of empirical literature on networks. We use social networks analysis tools to map the social ties of criminals in the US Mafia. To what extend is the Mafia a randomly generated network? Do criminals connect to other criminals that are similar to them, and similar with respect to what? What characteristics contribute to the formation of a leaders, *key players*, i.e. the most central agents in the network? What kind of network formation can rationalize our findings?

We also hope that the U.S. experience with Cosa Nostra and this wealth of information may offer clues to promising control techniques in countries where organized crime is on the rise (Jacobs and Gouldin, 1999).

2 The Origin of American Mafia

Before presenting our empirical study it is important to contextualize this study. We do this proving some background about when these "made" men came to the U.S. and how the Mafia operated in the 1960s when the FBN was filing the records we analyze.

Historians define two waves of immigration from Sicily, before and after World War I (WWI). Before WWI immigrants were mainly driven by economic needs. Several Mafia bosses, like Lucky Luciano, Tommaso Lucchese, Vito Genovese, Frank Costello, etc, were children of these early immigrants. Even though between 1901 and 1913 almost a quarter

of Sicily's population departed for America, many of these early immigrant families were not from Sicily. During those years around 2 million Italians, mainly from the south emigrated to the U.S. (Critchley, 2009). These baby immigrants later became street gang members in the slums; they spoke little Italian, and worked side by side with criminals from other ethnicities, mainly Jews and Irish (Lupo, 2009).

Lured by the criminal successes of the first wave of immigrants, and, quite paradoxically, facilitated by prohibitionism, the second wave of immigrants that later became Mafia bosses were already criminals by the time they entered the U.S.. Charles Gambino, Joe Profaci, Joe Bonanno, and others were in their 20s and 30s when they first entered the U.S., and they were all coming from Sicily.[6] Another reason for this selection of immigrants was the fascist crack-down of the Mafia, which forced some of these criminals to leave Sicily. After the second wave of immigration the Mafia became more closely linked to the Sicilian Mafia and started adopting its code of honor and its tradition.[7]

In 1930 and 1931 these new arrivals led to a Mafia war, called the Castellamare war, named after a small city in Sicily where many of the new Mafia bosses were coming from. The war lasted until Maranzano, who was trying to become the "Boss of the Bosses," was probably killed by Lucky Luciano, who had joined the Masseria Family.[8] This war put Lucky Luciano at the top of the Mafia organization but led also to a reaction by the media and by the prosecutors. In 1936 Thomas E. Dewey, appointed New York City special prosecutor to crack down on the rackets, managed to obtain Luciano's conviction with charges on multiple counts of compulsory prostitution. Luciano served only 10 years of the 30 to 50 years sentenced. In 1946 thanks to an alleged involvement in the Allied troops' landing in Sicily he was deported to Italy, from where he tried to

[6]Bandiera (2003) analyzes the origins of the Sicilian Mafia, highlighting how land fragmentation, missing rule of law, and predatory attacks generated demand for private protection.

[7]See Gosch and Hammer (1975)

[8]Previously, in order to end the power-struggle between Masseria and Maranzano Lucky Luciano had offered to eliminate Joe "the Boss" Masseria, which he did at an Italian restaurant, feeding him with lead.

keep on organizing "the organization." Between 1950 and 1951, the Kefauver Committee, officially the Senate Special Committee to Investigate Crime in Interstate Commerce, had a profound impact on the American public. It was the first committee organized to gain a better understanding of how to fight organized crime, and the main source of information was a list of 800 suspected criminals submitted by FBN's Commissioner Anslinger, most likely an early version of our records (McWilliams, 1990, pg. 141). But the Committee could not prove the existence of a Mafia.

After Luciano's expatriation several other Families headed the organization: Costello, Profaci, Bonanno, and Gambino. Family ties were of uttermost importance. According to Bonanno's autobiography (Bonanno, 1983), he became the Boss of the Bosses in part by organizing in 1956 the marriage between his son Bill and the daughter of Profaci, Rosalia. In 1957 Gambino took over the leadership.

Throughout the 1950s the FBN continued to investigate the Mafia, but until in 1957 an unexpected event happened the Mafia was not under Media's attention. The event was the lucky raid of an American Mafia summit, the "Apalachin meeting." Police raided the meeting and over 60 underworld bosses were detained. Joe Bonanno managed to flee, and was later known as Joe "Bananas." After that meeting everybody had to agree with the FBN's view that there was one big, and well organized Mafia. This meant probably the beginning of the end of the American Mafia. Robert Kennedy, attorney general of the United States, and even J. Edgar Hoover, head of the Federal Bureau of Investigations, joined Harry J. Anslinger, the U.S. Commissioner of Narcotics, in his war against the mob. The same years a permanent Senate Select Committee was formed, the McClellan commission. Anslinger's FBN conducted the investigative work and coordinated nationwide arrests of Apalachin defendants.

Lucky Luciano died of a heart attack at the airport of Naples in 1962. In the 50s and 60s *Cosa Nostra* was governed by a *Commissione* of from 7 to 12 bosses, which also acted

as the final arbiter on disputes between Families. The remaining 10 to 15 families were smaller and not part of *Cosa Nostra*'s governing body. The total estimated number of members was around 5,000 (Maas, 1968). The 800 profiles that we have access to thus do not cover all members, but certainly the most important ones.

After learning that had been marked for execution Joe Valachi became the first and most important informer for the FBN and later the FBI starting in 1962,[9] and revealed that the *Cosa Nostra* was made of approximately 25 Families. Each Family was structured in hierarchies with a boss, *Capo Famiglia*, at the top, a second in command, called underboss, *Sottocapo*, a counselor, *Consigliere*, and several capo, *Caporegime*, captains who head a group of soldiers (*regime*) (Maas, 1968). Figure 2 shows how in 1963, thanks to Valachi's testimony, a U.S. Senate commission set up to investigate organized crime, called the McClellan commission, drew the Bonanno Family "tree" structure. And Valachi provided information on many more families. Since our data represents a snapshot of what the authorities knew in 1960 they do not contain information about the Family each member belongs to. Nevertheless, we'll see that the pattern of connections is clearly informative about the structure of *Cosa Nostra*. Joe Valachi's testimony confirmed the FBN's view (which at the time wasn't the FBI's one) that the Mafia had a pyramidal structure with connections leading toward every single member (in our data almost the whole network is connected and the average path length is just 3.7).

3 Murder Incorporated

The Mafia, and more generally networks of criminals or gangs have tremendous influence on crime. They amplify delinquent behavior. In the sociological literature, this is referred as the social facilitation model, where gang members are intrinsically no different from

[9] Jacobs and Gouldin (1999) provide a relatively short overview about law enforcement's unprecedented attack on Italian organized crime families following Valachi's hearings.

no-gang members in terms of delinquency propensity. If they do join a gang, however, the normative structure and group processes of the gang (network) are likely to bring about higher rates of delinquency. Gang membership is thus viewed as a major cause of deviant behavior (see, e.g. Thornberry et al., 1993).[10] Moreover, crimes like racketeering, drug trafficking, gambling, bootlegging, etc require a group to be properly "organized."[11]

Empirical evidence on the relationship between crime and typically unobserved criminal organizations is, however, scarce, to say the least. The information provided in our data allows us to collect evidence on the presence of *Cosa Nostra* across cities. We investigate to what extent the presence and the density of the Mafia is correlated with local levels of violent (murder, rape, robbery, and assault) and non-violent (burglary, larceny, and auto theft) crimes. Crime rates are based on city-level uniform crime reports.[12]. Table 1 shows that the cities included in our data set, i.e. those where the Mafia operates, ("Mafia cities," hereafter) are the US cities where more violent crimes are committed. The variance of these crimes are higher in Mafia cities which is indicative of the *modus operandi* of the Mafia organization where periods of relative "peace" among different families alternate with "Mafia wars," with peak crime levels. Most of these cities are also cited by Valachi to be the ones where Families were active (Maas, 1968).

Among the Mafia cities the top panels of Figure 3 shows the raw correlation between the number of Mafia members per 100,000 inhabitants and local levels of violent and non-violent crimes, and in the bottom panel the one between the Mafia network structure and again violent and non-violent crimes. The network structure is measured by the average eigenvector centrality of the Mafia members that live in the city.[13] It appears

[10] In his very influential theory of differential associations, (Sutherland, 1947) emphasizes that the role of gangs is to facilitate the learning of crime technology.

[11] See Alexander (1997) for an economic analysis of racketeering with collusion.

[12] We use the same data constructed by Levitt (1997), which consist of a panel of roughly 60 US cities, with observations running from 1970 to 1992. The data are freely downloadable from the Internet at http://emlab.berkeley.edu/replications/mccrary/index.html

[13] The eigenvector centrality index assigns larger scores to nodes that are linked to high-scoring nodes than to nodes that are linked to low-scoring nodes. In appendix B we discuss this index in more detail. Google's page rank is a variant of the eigenvector centrality index (Bonacich, 1987).

that while the estimated correlations between Mafia indicators and crimes are only mild for non-violent crimes, they are clearly strong and positive for violent crimes.

Even though we are comparing the presence of the Mafia in 1960 with the average crime levels between 1970 and 1992, with persistent crime rates such correlation would also emerge if Families would systematically choose to reside in cities that were more appropriate for racketeering and that happened to have higher crime rates. In Figure 4 we address the selection issue plotting the city fixed effects within a typical crime regression that controls for standard determinants of the local crime rates (including income per capita and growth, police force, percentage of blacks, government spending in education and welfare, population age structure) on the presence and the density of the Mafia. The positive correlation between the Mafia structure and violent crime rates persists, while non-violent crimes do not seem to be positively related to the Mafia.

While we cannot claim that selection is not part of the story, it is unlikely to be the only part. Cracking down on Mafia and its violent culture is likely to lower violent crime rates. This happened in the US once the government started to severely prosecute mobsters, and might still happen in the Southern parts of Italy where the Mafia is still very active. We claim that analyzing the network within *Cosa Nostra* helps prosecutors to understand how *Cosa Nostra* works and how to best fight it.

4 Descriptive Evidence About *Cosa Nostra* and Its Members

4.1 The FBN Records

The criminal files come from an exact facsimile of a huge Federal Bureau of Narcotics report of which fifty copies were circulated within the Bureau starting in the 1950s. These files come from more than 20 years of investigations, and several successful infiltrations

by undercover agents (McWilliams, 1990). Given that in the U.S. there were an estimated 5,000 members active during those years the list represents a certainly non-random sample of *Cosa Nostra* members. More active and more connected mobsters were certainly more likely to be noticed and tracked, which is probably why most, if not all, big bosses that were alive at the time have a file. Later we will address how a random and a non-random selection influence our analyses.

4.2 Individual Characteristics of the Members

Before analyzing how criminals are connected within the Mafia it is instructive to describe the members based on the information contained in the criminal records. Let us first date the data. Given that the distribution of the year of first arrest has basically full support within the range 1908-1960 (the only year without a first arrest is 1910) one can infer that the data refer to what the authorities knew in 1960.[14] The records do not report any death, thus don't include big bosses that were killed before 1960, i.e. Albert Anastasia boss of one of the 5 New York City families, the Gambino one.[15]

Table 2 shows that the average year of birth is 1911, thus the average age is 49 years. The youngest member is 23 years old, while the oldest one is 60. Half of the mobsters reside in either New York, or in New Jersey, and probably entered the U.S. through Ellis Island. Indeed, 29 percent were born in Sicily and another 10 percent in other regions of Italy. Most remaining mobsters were born in the United States but were of Italian origin as it represented a prerequisite to become a member. 75 percent of members are married, but only 60 percent of these are reported to have children (0.44/0.75). The overall average number of children is 1 and is 2.14 among members with children, equally divided between sons and daughters. 18 percent of members are married to someone who

[14] Additional evidence is the following description in Michael Russo's file: "Recently (1960) perjured himself before a Grand Jury in an attempt to protect another Mafia member and narcotic trafficker."

[15] His brother Anthony "Tough Tony," instead, was killed in 1963 and is in our records.

shares her maiden name with some other member (*Connected wife*). These marriages are presumably endogamous within the Mafia. Observe that we are understating the percentage of marriages within the Mafia as some Mafia surnames might be missing in our data.[16] The FBN reports an average of 1.96 siblings per member, while the average number of recorded members that share the same surname is 1.62.

The average height is 5.6 feet, the average weight is 176 pounds.[17] Their criminal career starts early. They are on average 23 years old when they end up in jail for the first time, and the majority has committed some violent crime. Only 16 percent do not have an arrest record. We don't know the total number of crimes committed by the mobsters but we know in how many different types of crime they have apparently been involved. This number varies between 0 and 9 and the average is 2.58. We also know in how many different legal businesses they have interest in. This number varies between 0 and 5 and is on average equal to 1.

Tables 3 and 4 show the list of legal and illegal activities that at least 5 percent of members were involved in. Most mobsters owned restaurants, drugstores or were otherwise involved with the supply of food. Real estate, casinos, car dealerships, and import export were also common businesses. Among the illegal activities the most common one is drug trafficking (43 percent), maybe also because the information was gathered by the FBN. Twenty-six percent of members were involved in robberies and 23 percent in murders. Weapon offences and assaults are also quite common. Overall, several members are involved in violent crimes which is in line with the correlations between Mafia density and violent crimes presented in Section 3. Some crimes that are typically associated with organized crime, like gambling, extortions, and liquor offences (during prohibition) are highly represented as well.

[16] While it is also possible that some women might have a Mafia surname without being linked to any Mafia family, this is very unlikely conditional on being married to a Mafia associate.

[17] As a note, 18 percent of the mobsters are obese and 58 percent overweight.

4.3 Individual Characteristics in Relationship to the Network

Each criminal record contains a list of criminal associates. Figure 1 shows, for example, that Joe Bonanno was associated with Luciano, Costello, Profaci, Corallo, Lucchese, and Galante. There is no evidence about how the FBN established such associations, and why they were restricted to be 5. Undirected connections are clearly more numerous, as mobsters can be listed as associates in several records. Hence, we define two mobsters to be connected whenever at least one mobster lists the other mobster's last name in his record.[18]

Table 5 shows the list of members with the 10 highest and 10 lowest number of direct connections.[19] The FBN records contain for each mobster a paragraph about the mobster's activities within *Cosa Nostra*. In order to extract information on the level of importance of these criminals the variable *Top* counts the number of times the words "boss," "highest," "most," "head," and "top" are cited and *High* the number of times "high," "influential," "important," "leader," "leading," "powerful," and "representing" are cited. The *Apalachin* variable indicates whether the mobster attended the important 1957 Mafia meeting in the Upstate New York. The last column represents the historically reconstructed position within the Mafia.

Criminals with many connections are more likely to be recognized as high-ranked members, and more likely to have attended the 1957 meeting. Several members in the top distribution of the number of connections are bosses, i.e. Salvatore Lucania, *alias* Lucky Luciano, Vito Genovese, Antoni Accardo, Joe and Joseph Profaci. Salvatore Santoro and Salvatore Vitale were instead underbosses of the Lucchese and Bonanno Family. Criminals with the lowest *degrees*, instead, are mostly soldiers.

The number of connections, called degree in network analysis, are thus one way to mea-

[18] In other words, we construct a symmetric adjacency matrix of undirected connections.

[19] Since connection are based on surnames members with the same surname will share the same *degree*. This introduces some noise but dealing with the large variation in first names, i.e. Antonio/Tony/Anthony, would introduce even more noise in graphing the network.

sure the importance of members, but over the past years social network theorists have proposed different centrality measures to account for the variability in network location across agents and there is not a systematic criterium to pick up the "right" centrality measure for each particular situation (Borgatti, 2003, Wasserman and Faust, 1994).[20] Before showing the distribution of some of these measures we briefly define them (see Appendix B for a more technical definition). Unlike *degree*, which weights every contact equally, the eigenvector index weights contacts according to their centralities. The index takes direct as well as indirect connections and thus the whole network into account.[21] The *closeness* index represents the inverse of the average distance between a node (a member) and all the other nodes, and is a good measure for how isolated members are. The *betweenness* index measure the number of times a node is on the shortest path between two randomly chosen nodes, and is a good measure for how the member's capacity to act like a bridge between clusters of members.

These measures of individual centrality allow us to consider different nuances in the definition of a Mafia leader. Our data also allow us to construct a qualitative indicator of importance, based on the union of the *Top* and *High* variables discussed in this section. Table 6 collects descriptive statistics about our various indicators of importance.[22] The non-standardized *degree* varies between 1 and 71, and Figure 5 shows that the corresponding density is positively skewed. The eigenvector index (centrality) has a density that is very similar to the one of *degree*, while the one of *closeness* is more symmetrically distributed, meaning that most mobsters are neither too isolated nor too close within the network. The density of *betweenness*, instead, shows that very few mobsters represents bridges between subsets of the network, most likely Families. Finally, fifty-five percent of the mobsters appear to be high-ranked, meaning that their descriptions includes at least

[20] See also Sparrow (1991) for a discussion on centrality indices in criminal networks.

[21] As firstly noted by Granovetter (1973), weak ties (i.e. friends of friends) are important source of information. See Patacchini and Zenou (2008) for the role of weak ties in explaining criminal activities.

[22] Our individual centrality measures have been normalized such that the maximum values are equal to 100.

a word that implies leadership.

The last variable to be summarized in Table 6 is the interaction index. The index measures the exposure to what Bonanno (1983) calls, in uppercase, "Tradition" or Hess (1973) calls, in lowercase, "mafia," or "mafia culture" to distinguish it from "Mafia" the organization. Figure 6 helps us explain the index. It shows the current distribution at the zip code level of the members' surnames in Italy's phone directory.[23] Each circle is proportional to the number of surnames present within each zip code. Not surprisingly many surnames show up in Sicily, in Naples, and in Calabria. Many of these surnames appear also in large cities that were subject to immigratory flows from the south, like Milan, Rome, and Turin. For each members' surname we computed the probability that it shares a randomly chosen zip code located in the South of Italy with other surnames from the list. To be more precise, the index for member i is equal to 100,000 times the sum across zip codes j of the fraction of surnames of member i present in zip codes j times the fraction of surnames of the other members $(-i)$ in the same zip code:

$$interaction_i = 100,000 \sum_j \frac{\#surname_{i,j}}{\sum_j \#surname_{i,j}} \frac{\#surname_{-i,j}}{\sum_j \#surname_{-i,j}}. \quad (1)$$

The advantage of this index is that we can computed it for all surnames while information about the Italian community of origin would only be available for those born in Italy. The average index is equal to 3.7 per 100,000. Ten percent of the times the index is zero, either because the zip codes do not overlap or because the surname is not in the phone directory.

4.3.1 Characteristics of the Mafia Network

As pointed out by Jackson and Rogers (2007), the distribution of *degree* in a socially generated network is more unequal than in random networks. We start our empirical investigation by providing some evidence on the extent to which such empirical regularity

[23] Unfortunately we could not find the distribution of surnames in 1960.

find support in our real world network.

In random networks the log-frequency is linear in log-*degree*. Figure 7 shows that this is not true for the Mafia network. Links are not random. The interesting question is then: what are the forces shaping the associations between criminals? Are criminals more likely to be associated with criminals who established similar illegal businesses or do they prefer to diversify the risk of detection keeping a lower profile and thus associate themselves with criminals who operate different kinds of businesses?

Our data show that criminals tend, indeed, to associate with like-types both in illegal and legal activities. Table 7 shows the extent to which criminals associate with other criminals in similar kind of business, in illegal and legal activities. The tables show positive and significant correlation, called assortativity, for almost all crimes and legal businesses detailed in our data.

Given the non-randomness of the *Cosa Nostra* network, before moving to the empirical analysis of the determinants of the network structure we develop, based on historical accounts, a simple model of Mafia connections. The aim of the model is to show how a simple model of connections can generate a positively skewed distribution of *degree* and some interesting comparative static results.

5 A Simple Model of Connections

In this Section we develop a simple model of hierarchical connections.[24] The boss or equivalently the Family decides the optimal number of soldiers (s) and captains (c), the intermediaries who lead a group of soldiers. The boss or the captains do not face restrictions on the supply of subordinates, as there is evidence of excess-supply of low ranks (Maas, 1968). The total (reduced form) benefits of the whole organization, $\int_0^t u_r(s,c)\,dr$, depend on the number of different Families that are active (t). Given that the period

[24]See Baccara and Bar-Isaac (2008) for a game-theoretical model that rationalizes hierarchies.

between the Castellamare war (1929-1931) and 1960 was a relatively peaceful time inside *Cosa Nostra*, with an established *Commissione* that was resolving the disputes Families had between them, we do not model competition or war with other Families (Lupo, 2009). Moreover, after wiping out competing gangs, mainly the Irish and the Jewish ones, the Italian Mafia had gained complete control of racketeering. During those years the Mafia was behaving like a stable cartel, with self-enforcing strict rules to keep the cartel running: markets were geographically segmented; Family memberships were held stable to keep relative power unchanged; the status quo of leadership within Families was endorsed and insurrections were repressed collectively (Bonanno, 1983). We also do not model the choice of the underboss as in Bonanno (1983)'s words "he is just a figurehead who represents the Family on various family matters."

The *Commissione* decides its optimal size t depending on its ability to keep the Families united. Each Family represents a new tree of connections for which the *Commissione* he has to pay a constant cost γ. The total utility $v(\cdot)$ depends of the difference between benefits and costs. For simplicity all these variables are assumed to be continuous.

$$\begin{aligned}\max_{t,s,c} v\left(t,s,c\right) &= \max_{t,s,c}\left[v\left(\int_0^t u_r\left(s,c\right)\right)dr-\gamma t\right] \\ &= \max_{t}\left[v\left(\int_0^t \max_{s,c} u_r\left(s,c\right)\right)dr-\gamma t\right]\end{aligned}$$

$$s.t.\ \ \gamma>1, s>0, c>0$$

Assuming, without loss of generality, symmetry across Families, $u(\cdot)=u_r(\cdot)$, the first order condition with respect to t is

$$v'\left[tu\left(s^*,c^*\right)\right]-\gamma=0$$

where s^* and c^* represent the optimal choices in the number of soldiers and captains (see Falcone and Padovani (1991, pg. 99) for a discussion about soldiers who are under direct control of their boss). If the marginal benefits are decreasing in t, i.e. because of increased complexity, a *Commissione* with larger marginal costs would also have fewer members. To solve for the optimal number of soldiers and captains we maximize the benefits of each tree, $argmax_{s,c} v\left(u\left(s,c\right)\right) = argmax_{s,c} u\left(s,c\right)$.

We specify the following functional form for the expected benefits of a tree c:

$$u(s,c) = (s + ck(1-\tau)q^{ck})p^s. \tag{2}$$

Given that "(T)he size of the group varies depending on group leader's following," we assume that size to be fixed and equal to k (Bonanno, 1983). The number k, and in general the tree developing under a captain could just be modeled in the same way we model the boss's connections, or could for simplicity just be the solution of a model without lieutenants:

$$k = argmax_s sq^s, \tag{3}$$

where q^s represents the probability of success of the captain's business with s soldiers when each soldier generates a payoff of 1.

The rationale is that when a soldier with direct connections to the boss gets caught the boss might loose the entire business. This happens with probability $1-p$. Thus with probability p^s he keeps the profits of the entire business tree if s soldiers work *directly* for him. Bonanno (1983) presents a perfect example of such a trade-off. In 1931, after Lucky Luciano's killing of Maranzano, Joe Bonanno, the new boss of the Castellamare faction in Brooklyn, was offered to directly participate in Lucky Luciano's stake in the NYC clothing industry. But he overcame his "temptation-that of making money easily

but heedlessly" in order to avoid "a practice with a high risk of discovery." "Unlike Charlie Lucky" he wanted to keep his "name clean and out of the public eye."[25]

Bosses might, therefore, prefer to delegate parts of the business to captains. The probability the captain's business survives, thus none of his soldiers gets caught and no investigation is successful, is q^{ck}. If the captain's business gets disrupted this does not spill over to the boss. This insulating mechanism that develops makes prosection of bosses extremely difficult (Critchley, 2009)[26]

The profits are a linear function of the number of soldiers s and a linear function of the number of captains c. But in accordance with the actual workings of the Mafia captains keep a fraction τ of the profits ck, which are proportional to the number of soldiers they employ k (Raab, 2006, pg.5). Business is structured like a franchised company, where in exchange of τ bosses provide protection.

From the first order conditions (shown in the appendix) one can show that the optimal number of soldiers and captains is

$$s^* = \frac{(1-\tau)\, q^{\frac{-1}{\ln q}}}{\ln q} - \frac{1}{\ln p}$$

and

$$c^* = \frac{-1}{k \ln q}.$$

[25]See Reuter (1985) for a discussion about the optimal size of criminal organization and, more generally, about providing incentives for loyalty.

[26]During Joe Valachi initiation rite boss Maranzano explained that if a soldier has to see his boss he has to ask his captain first. "A soldier ain't allowed to go running all the times to his boss" (Maas, 1968).

The comparative static results for soldiers and captains are

$$\begin{aligned}
\frac{\partial s^*}{\partial p} &= \frac{1}{(\ln p)^2}\frac{1}{p} > 0, \\
\frac{\partial s^*}{\partial \tau} &= \frac{-q^{\frac{-1}{\ln q}}}{\ln q} > 0, \\
\frac{\partial s^*}{\partial q} &= \frac{\ln q\,(1-\tau)\left(\frac{1}{qq^{\frac{1}{\ln q}}\ln q} - \frac{1}{q^{\frac{1}{\ln q}+1}\ln q}\right) - \frac{1}{q}(1-\tau)\,q^{\frac{-1}{\ln q}}}{(\ln q)^2} \\
&= \underbrace{\frac{1}{\ln q}(\tau-1)\left(\frac{1}{q^{\frac{1}{\ln q}+1}\ln q} - \frac{1}{qq^{\frac{1}{\ln q}}\ln q}\right)}_{<0} + \underbrace{\frac{1}{qq^{\frac{1}{\ln q}}\ln^2 q}(\tau-1)}_{<0}, \quad <0
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial c^*}{\partial k} &= \frac{1}{k^2 \ln q} < 0 \\
\frac{\partial c^*}{\partial q} &= \frac{1}{(k\ln q)^2} k \frac{1}{q} > 0.
\end{aligned}$$

The results are intuitive and would hardly change if we made the model more complicated, i.e. introducing a competitive labor market for affiliates, or a competitive environment between Families.

The number of direct soldiers increase with the probability of success p, with the "tax" paid to captains, and go down with the probability of success of captains q. The number of captains, instead, does not depend on the probability of success of soldiers p, but does depend positively on their own probability of success and negatively on the number of their crew.

Given the closed form solutions for s^* and c^* one can easily simulate the distribution of total connections of the boss, $s + c$, for different levels of p and q. Figure 8 shows that even if we draw p and q from a uniform distribution the model does not predict a

uniform distribution of connections, but rather a positively skewed one, for both soldiers and captains. This particular shape does not depend on τ, or on the range of p and q, or on their difference.

The number of affiliates depends on the boss's p and q, the ability to prevent soldiers and captains to become informers for the FBI or the FBN. What do these abilities depend on? Fear of retaliation for breaking the vow of silence, called *omertà*, has certainly been a key factor for the success of the *Cosa Nostra*. Some members were well known for the violence, their bloodshedding attitude, among these is Albert Anastasia also known as the "Mad Hatter" and "Lord High Executioner," killed in 1957 in a barber shop because of an internal war between Vito Genovese and Frank Costello. Social ties, in particular, family ties between the boss and his affiliates are also likely to keep p and q high. Bosses with larger families should therefore have more connections. Moreover, captains with family ties might also have a lower intermediary cost τ. In the next Section we present the first empirical test of these hypotheses.

6 Which Characteristics Predict the Importance of a Mafia Member?

The following analysis will try to highlight the forces that might influence p and q, and, therefore, the number of connections mobsters have. Since the data do not allow us to distinguish soldiers from captains, the equation that we estimate can be interpreted as a linear approximation of $c^* + s^*$. While for the sake of simplicity the model did not take into account the quality of the links, in the previous Sections we saw that there are several ways to measure how central, or how connected a member is within the network. We start our regression analysis using the eigenvector index as dependent variable, as it depends both on the number and the quality of its connections, on the whole sample of

mobsters. But several of them are likely to be soldiers, and thus to be more coherent with the theoretical model later we show how our results change when we restrict the analysis to top mobsters.

6.1 Descriptive Evidence

Before moving to the regressions a graphical analysis can shed light on some organizational rules of the Mafia. Figure 9 shows that all centrality measures grow steadily with age. This finding confirms Ianni and Reuss-Ianni (1972, pg.130)'s anthropologic results about the importance of age in determining the leadership positions. Unlike typical income profiles there is absolutely no evidence about a inverse U-shaped relationship between centrality and age. The difference between the minimum age (23) and the maximum age (60) doubles the degree, triplicates the eigenvector index, and more than triplicates *betweenness*. Closeness, instead, shows a very steep increase up to age 30 and than flattens out, but the overall increase is more modest.

Figure 10 shows how the interaction index, our continuous proxy for "mafia culture" influences centrality. The overall patterns are less clear than for age, but when interactions are very high all centrality measures appear to increase as well. Figure 11 shows that eigenvector centrality is low in the absence of daughters or sone, but that sons and daughters seem to be equally valuable in increasing the father's centrality. We will say more about this later when we analyze the importance of intermarriage. Figure 12 shows that the number of types of crimes and of businesses do also positively influence mobster's centrality within the network, especially up to the number 4. Above 4 types of crimes allegedly committed the evidence is more noisy as there fewer observations are available. We present how categorical variables influence network centrality in the next Section, where we move to a regression framework.

6.2 The Eigenvalue Index

In Table 8 we use the eigenvector centrality measure as our dependent variable, though more generally we assume the following linear relationship between the k-th index of importance c_i^k of individual i and his observable characteristics X_i: $c_i^k = \beta' X_i + e_i$, $k = 1, ..., 4$.[27]

Each regression controls for the number of mobsters that in our data share the same surname. We call this variable "Extended family members" despite the possibility that some of these mobsters might not be related to each other. Given that our connections are based on surnames we added this variable is to control for mechanical effect that an increasing number of mobsters might have on our centrality measures. The coefficient is always around 5, as 5 is also the number of associates contained in each criminal file.

In the first column we regress the eigenvector index on individual characteristics that are plausibly exogenous. Each additional year of age increases the index by 0.247 (2 percent), which is a very large effect and, as we mentioned, is consistent with Ianni and Reuss-Ianni (1972)'s account of hierarchies based on generations. Another variable that has a significant positive influence on centrality (20 percent) is migrating from Sicily. Not only was the US Mafia an Italian enterprise, it valued direct links with Sicilians more than with people from other parts of Italy. Sicilian kin-centered social system, with its code of honor and vow of silence, forms the building block for the Mafia. Bonanno would write that among Italians he felt safe only around Sicilians. Moreover, nativity does not fully capture adherence to the mafia code of law as 60 percent of the mobsters were born in the U.S. But the interaction index, which depends on the geographical distribution of last names, and proxies for community ties and exposure to the "Tradition," has also a positive and significant effect on network centrality. Members from Sicily and those with stronger ties to the Mafia culture could probably be trusted more, as they were more likely to adhere

[27] See Appendix for a detailed description of the different indices.

to the *omertà*, increasing p and q. Height and weight of the mobster does not influence his status.

In column 2 we control for the nuclear family structure of the mobster. Each additional child increases the eigenvector index by one, while it doesn't matter whether the child is male or female. This finding challenges the criminological view that within the Mafia male children are more valuable than female ones because they represents potential "workforce." Our leading explanation for this finding is that in a male only society like the Mafia "connected" girls (probably in excess demand) could be married strategically.

The number of siblings, and being married or divorced does not influence the centrality index. This is in line with Falcone and Padovani (1991, pg. 113)'s view that unlike the Mafia in Italy the American Mafia adapted a more liberal view toward divorce.

Being married to a wife who is connected, instead, leads to an almost 50 percent increase in the index. These findings, again, seem to suggest that higher p and q, due to more trusted links, increase the optimal number of connections. We cannot rule out the possibility that more connected mobsters are also more likely to find, or be given, a connected wife. While these alternative interpretations do not matter when the only purpose is to discover the leading figures within the Mafia, one has to be careful in giving a causal interpretation to these estimates. In the third column we control for the known legal and illegal activities of the mobster. The leading mobsters are active in New York (+37.5 percent) and New Jersey (+51 percent). Criminals who were known to have committed violent crimes have an index that is 15 percent larger (+1.863). If being violent reduces the risk of betrayal this finding is consistent with our model. An arrest increases the importance by 2.6 (20 percent), though, again, higher ranked individuals might be more likely to be arrested. Age at first arrest is negative, showing that not only age, but also experience increases the level of centrality in the network (young members are often recruited in jail; Ianni and Reuss-Ianni, 1972, pg. 45).

The remaining two variables measure how many different types of crime and different types of businesses the mobster was involved in. Both variables are positive. This is consistent with more able criminals being more able to diversify risk. But the number of business types (+16 percent) are a better predictor for "key players" than the number of crime types (+4 percent). The last column combines all the regressions in one without notable changes. The combined regressors explain almost 30 percent of the variability of the eigenvector index.

6.3 Other Network Centrality Indices

Eigenvector centrality represents only one way to measure centrality in the network. Other indices capture different nuances of centrality. In Table 9 we look at different measures of centrality and importance using the same specification used in the last column of Table 8. In order to highlight how different measures are able to capture different characteristics of the network, and since there is no systematic way to decompose the different centrality measures based on the importance of direct and indirect links, we developed a simple statistical way to accomplish the same goal: we simply add as a dependent variable the residual (ϵ_i^j) of a linear projection of the alternative measures of centrality (c_i^j; $j \neq 1$) on *degree* (c_i^1): $c_i^j = \alpha + \beta c_i^1 + \epsilon_i^j$; $j \neq 1$.

The coefficients on the residuals of a particular centrality index measure the nuances captured by that particular index with respect to a simple count of direct connections. Column 2 replicates the last column of Table 8 to eases the comparison with the other measures. Comparing columns 1 and 2 shows that the coefficients are not very different when *degree* is used to proxy for leadership. The only coefficients that seem to be smaller are the ones on residing in New York or New Jersey, and the one on the types of businesses. Our simple test for the significance of these differences that looks at the residuals confirms that these differences are significant. This means that living in New York and New Jersey,

and the number of businesses increase the eigenvector centrality not only through the direct links but also through the indirect ones, while for all the other variables that are significant in column 2 but not in column 3 only direct links matter.

Closeness, an inverse measure of the average distance from the other members, and *betweenness*, the ability to build bridges, capture different aspects in the definition of individual importance or power within a network. Column 4 shows that overall the coefficients for *closeness* do not differ substantially from those seen before. The main difference is that height reduces the average distance between members, while the types of crime have no significant effect on how close a member is to the other mobsters. *Closeness* is also more than any other index dependent on kinship. The "Sicily" coefficient, the interaction index coefficient, and the "Connected" wife coefficient are all highly significant. Moreover, column 5 shows that a large part of these effects are driven by the indirect links (cannot be explained by *degree* alone).

Results that use *betweenness* as a dependent variable shown in column 6 present very interesting findings: there is no evidence that bridges are build through marriages. The coefficient on the *Connected wife* dummy is precisely estimated to be close to zero. Strategic marriages are thus confined to happen within "friendly" clans, and not across clans that wouldn't otherwise be connected. The number of different crimes committed have also no bridging capacity, while businesses do.

Experience and age, instead, influence *betweenness* beyond the increase in *degree*. And so does being Sicilian, residing in New York or New Jersey, and being active in several businesses.

6.4 Qualitative Measures of Importance

In column 2 of Table 10 we exploit qualitative information on the mobsters contained in the records. In particular, we use the same specification as before but the dependent

variable counts the number of times the following words are used in describing a mobster: "boss," "highest," "most," "head," 'top," "high," "influential," "important," "leader," "leading," "powerful," and "representing." Overall the results confirm the importance of directly migrating from Sicily, being old and experienced, residing in New Jersey or New York, being active in many types of businesses, and using violence. But marrying a connected wife, having many siblings, having strong community ties (interactions), and the number of crimes committed are not anymore significantly related to leadership. This lack of significance leads to a lower R-squared (8.2 versus 28.7 percent). While there might be superior ways to extract the information know by the FBN (we did try several qualitative measures), network based measures seem to capture additional attributes of leadership.

6.5 *Indegree*, Random and Non-Random Truncation of the Sample

Table 11 shows two types of robustness checks. In the first instead of using undirected links we use only the *indegree*, meaning the number of times someone appears as an associate in criminal profiles of other mobsters. Given that *outdegree* is bounded at 5 it is not surprising that the coefficients are driven by *indegree*, and thus are very close to the one that correspond to the undirected *degree* shown in the first column of Table 9. The second and more important robustness check simulates random truncations of our sample. Since our initial sample is likely to be truncated as well—some scholars estimate that there as many as 5,000 "made" men across the U.S—, we try to address the possible bias that such a truncation might generate. Columns 3 and 4 show the mean and the median coefficients obtained in regressions based on 500 different samples where we randomly truncated 50 percent of the observations. *Indegree* is based on 50 percent of the sample only and, as before, standardized to lie between 0 and 100. There is no

evidence that randomly truncating the sample biases the coefficients.

Our sample, though, is certainly a non-random sample of mobsters. Those with more connections and more power are certainly more likely to be noticed by law enforcement agents. To address the effect of non-random truncation in Table 10 we separate the sample depending on whether the eigenvector index is above or below median. The theoretical model refers to leaders who select their crew. The determinants that we argued influenced p and q should, therefore, be stronger predictor of leadership among the top-ranked mobsters than among the lower-ranked ones. In column 3 and 4 we split the sample into mobsters that are above or below the median eigenvector index. Violence, types of legal and illegal activities, a connected wife, more family members are indeed strong predictors of leadership only among the top leaders. And, while not significant, the coefficients on year of birth, interaction index, and number of children are large for top leaders and close to zero otherwise. The only seemingly puzzling result is the lack of significance for being born in Sicily. But only 76 out of 229 Sicilians are lower-ranked mobsters. The likelihood of being top-ranked is 67 percent for Sicilians and only 43 percent for non-Sicilians. Sicilian mobsters are more likely to be in the upper distribution of leadership, but once there they are above the median, they are not more likely to be on top. For lower-ranked criminals, instead, residing in New Jersey or New York allows them to establish more connections, and so does jail time. All the other coefficients are not significantly different from zero.

7 Concluding remarks

This paper presents the first thorough micro-level analysis of the US Mafia network. Beside testing sociological and historical views about the functioning of these criminal networks, we develop a simple model of connections that highlights the trade-off of connections: they increase profits and power but also the risk of detection. More connections

mean more potential informers. Family ties, violence, and exposure to Mafia culture reduce the probability of defection and increase the number of connections. Unlike economic organizations hierarchies depend crucially on kinship. Our results highlight how intermarriage shapes the network. Women are used to foster the Family's network centrality, but only within trusted Families. Woman are not used to bridge Families that are not otherwise closely connected. Trust shapes the network. Where values are shared and the mafia culture is strong connections are more stable and thus more numerous. Coherent with mafia culture, how central members are within the network increases steadily with age. And more central members are going to have more businesses, legal and illegal ones.

Mathematical modeling necessary to tie explicit social network structure back to the diffusion of economic behavior is expanding at a radical pace. Empirical evidence on observed social network phenomena can thus be helpful to inform such a modeling. But empirical evidence has also an important value per se, for policy purposes. If social connections are the driving forces of the phenomenon under consideration and if their structure is non-random, as in our case, a detailed study of the characteristics of the network might reveal some relevant features of social structure that can guide crime preventing policies. A targeted policy identifying "key players" in a given area may be an effective way to reduce crime (Ballester et al., 2006). A key player is an individual belonging to a network of criminals who, once removed, leads to the highest aggregate delinquency reduction. In practice, the planner may want to identify optimal network targets to concentrate (scarce) investigatory resources on some particular individuals, or to isolate them from the rest of the group, either through leniency programs, social assistance programs, or incarceration. The success of such policies depends on the ability to identify a social network structure. Only detailed empirical studies on real-world social networks can provide guidance in this direction.

A The Model

The first order conditions with respect to s and c are:

$$\begin{aligned} s &: \; p^s + \ln p \left(s + ck\left(1-\tau\right) q^{ck}\right) p^s = 0 \\ c &: \; k\left(1-\tau\right) q^{kc} p^s + k \ln qck \left(1-\tau\right) q^{kc} p^s = k\left(1-\tau\right) q^{kc} p^s \left(1 + \ln qck\right) = 0 \end{aligned}$$

The second order conditions with respect to s and c are:

$$\begin{aligned} ss &: \; \ln pp^s + \ln pp^s + \ln p \ln p \left(s + ck\left(1-\tau\right) q^{kc}\right) p^s \\ &= \; 2 \ln pp^s < 0 \\ cc &: \; k^2\left(1-\tau\right) q^{kc} p^s \ln q < 0 \\ sc &: \; \left(\left(1-\tau\right) q^{kc} k + k^2 \ln qc \left(1-\tau\right) q^{kc}\right) p^s \ln p \\ &= \; 0 \end{aligned}$$

B Centrality Measures

Let $N = \{1, \ldots, n\}$ be a finite set of agents in network $\mathbf{g}$. Let us define $\mathbf{G}$ the $n-$square adjacency matrix of he network $\mathbf{g}$, i.e. the matrix that keeps track of the direct connections in this network, where $g_{ij} = 1$ if i and j are directly linked, and $g_{ij} = 0$, otherwise.[28]

The simplest index of connectivity is the number of direct links stemming from each agent i in the network, i.e. degree centrality:

$$c_i^1 = \sum_{j=1}^{n} g_{ij}.$$

The definition of centrality is thus based on the number of direct links only. A variant of

[28]Measurement errors in the definition of the actual connections are clearly unavoidable in our study. Some names are blanked out and many lower-ranked mafiosi were not filed.

simple *degree* is eigenvector centrality, which also takes into consideration indirect links:

$$c_i^2 = \lambda^{-1} \sum_{j=1}^{n} g_{ij} c_j^2$$

where λ is the highest eigenvalue of matrix $\mathbf{G}$. The formula implies (recursively) that the centrality of individual i is proportional to the sum of centralities of the individuals she/he is connected to. It thus can be high even if she/he has low *degree*.

The standard measure of *closeness* centrality of individual i is given by:

$$c_i^3 = \frac{1}{\sum_{j=1}^{n} d_{ij}}$$

where d_{ij} is the geodesic distance (length of the shortest path) between individuals i and j.[29] As a result, the *closeness* centrality of individual i is the inverse of the sum of geodesic distances from i to the $n-1$ other individuals and can be regarded as a measure of how long it will take information to spread from a given member to other members in the network.

Betweenness indexes derive from the number of optimal paths across (or from) every node. It can be defined as:

$$c_i^4 = \sum_{j,l}^{n} \frac{a_{jl,i}}{a_{jl}}$$

where j and l denote two given agents in $\mathbf{g}$, $a_{jl,i}$ is the number of shortest paths between j and l through i, and a_{jl} is the number of shortest paths between j and l.

[29] The length of a shortest path is the smallest k such that there is at least one path of length k from i to j. We can identify such a length by computing $\mathbf{G}$, $\mathbf{G}^2$, $\mathbf{G}^3$, ..., until we find the first k such that the (i,j)th entry of $\mathbf{G}^k$ is not zero.

References

MAFIA: The Government's Secret File on Organized Crime. By the United States Treasury Department, Bureau of Narcotics. HarperCollins Publishers, 2007.

Howard Abadinsky. *Organized Crime*. Wadsworth Publishing, 1990.

Barbara Alexander. The Rational Racketeer: Pasta Protection in Depression Era Chicago. *The Journal of Law and Economics*, 40(1):175–202, 1997.

Mariagiovanni Baccara and Heski Bar-Isaac. How to Organize Crime. *Review of Economic Studies*, 75(4):1039–1067, 2008.

Coralio Ballester, Antoni Calvó-Armengol, and Yves Zenou. Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417, 09 2006.

Oriana Bandiera. Land reform, the market for protection, and the origins of the Sicilian mafia: theory and evidence. *Journal of Law, Economics, and Organization*, 19(1):218, 2003.

Marianne Bertrand and Antoinette Schoar. The role of family in family firms. *Journal of Economic Perspectives*, 20(2):73–96, Spring 2006.

Phil Bonacich. Power and Centrality: A Family of Measures. *American Journal of Sociology*, pages 1170–1182, 1987.

Joseph Bonanno. *A Man of Honor: The Autobiography of Joseph Bonanno*. St. Martin's, 1983.

Steve P. Borgatti. *The Key Player Problem*, pages 241–252. R. Breiger, K. Carley and P. Pattison, Committee on Human Factors, National Research Council, 2003.

Yann Bramoullé and Bernard Fortin. *The Econometrics of Social Networks*. New Palgrave Dictionary of Economics, steven n. durlauf; lawrence e. blume edition, 2010.

David Critchley. *The Origin of Organized Crime in America: The New York City Mafia, 1891–1931*. Routledge, 2009.

Giovanni Falcone and Marcelle Padovani. *Cose di Cosa Nostra*. Rizzoli, 1991.

Gianluca Fiorentini and Sam Peltzman. *The Economics of Organised Crime*. Cambridge University Press, 1997.

Diego Gambetta. *The Sicilian Mafia: the business of private protection*. Harvard Univ Press, 1996.

Nuno Garoupa. Optimal law enforcement and criminal organization. *Journal of Economic Behavior & Organization*, 63(3):461–474, July 2007. URL `http://ideas.repec.org/a/eee/jeborg/v63y2007i3p461-474.html`.

Martin A. Gosch and Richard Hammer. *The Last Testament of Lucky Luciano*. Little, Brown, 1975.

Mark S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78: 1360–1380., 1973.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Does Culture Affect Economic Outcomes? *The Journal of Economic Perspectives*, 20(2):23–48, 2006.

Henner Hess. *Mafia and Mafiosi: The structure of power*. Ashgate, 1973.

Francis A.J. Ianni and Elisabeth Reuss-Ianni. *A Family Business: Kinship and Social Control in Organized Crime*. Russell Sage Foundation, 1972.

Matthew O. Jackson and Brian W. Rogers. Meeting strangers and friends of friends: How random are socially networks? *American Economic Review*, 97:890–915., 2007.

James B. Jacobs and Lauryn P. Gouldin. Cosa nostra: The final chapter? *Crime and Justice*, 25:pp. 129–189, 1999. ISSN 01923234. URL http://www.jstor.org/stable/1147609.

Vimal Kumar and Stergios Skaperdas. *Criminal Law and Economics*, chapter On The Economics of Organized Crime. Edward Elgar Pub, 2009.

Steven D. Levitt. Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review*, 87(3):270–90, 1997.

Steven D. Levitt and Sudhir Alladi Venkatesh. An economic analysis of a drug-selling gang's finances. *The Quarterly Journal of Economics*, 115(3):755–789, August 2000.

Salvatore Lupo. *Quando la Mafia Trovò l'America*. Einaudi, 2009.

Peter Maas. *The Valachi Papers*. Putnam, New York, 1968.

John C. McWilliams. *The Protectors: Harry J. Anslinger and the Federal Bureau of Narcotics, 1930-1962*. Univ of Delaware Press, 1990.

Carlo Morselli. Career opportunities and network-based privileges in the Cosa Nostra. *Crime, Law and Social Change*, 39(4):383–418, 2003.

Eleonora Patacchini and Yves Zenou. The strength of weak ties in crime. *European Economic Review*, 52:209–236, 2008.

Alejandro Portes. Social Capital: Its Origins and Applications in Modern Sociology. 24: 1–24, 1998.

Selwyn Raab. *Five Families: the Rise, Decline, and Resurgence of America's Most Powerful Mafia Empires*. St. Martin's Griffin, 2006.

Peter Reuter. *Disorganized Crime: the Economics of the Visible Hand*. MIT press, 1983.

Peter Reuter. The organization of illegal markets: An economic analysis. 1985.

Peter Reuter. *Handbook of organized crime in the United States*, chapter Research on American Organized Crime. Greenwood Press, Westport, CT, 1994.

Friedrich Schneider and Dominik H. Enste. Shadow economies: Size, causes, and consequences. *Journal of Economic Literature*, 38(1):77–114, March 2000.

Malcolm K. Sparrow. The Application of Network Analysis to Criminal Intelligence: an Assessment of the Prospects. *Social Networks*, 13(3):251 – 274, 1991. URL http://www.sciencedirect.com/science/article/B6VD1-466DRKY-14/2/4785fa0af47facaa260a47

Edwin H. Sutherland. *Principles of Criminology*. Chicago: J.B. Lippincott, 1947.

Terence P. Thornberry, Marvin D. Krohn, Alan J. Lizotte, and Deborah Chard-Wierschem. The Role of Juvenile Gangs in Facilitating Delinquent Behavior. *Journal of Research in Crime and Delinquency*, 30:55–87, 1993.

John R. Wagley. Transnational Organized Crime: Principal Threats and US Responses. Congressional Research Service Report for Congress RL33335, March 2006.

Stanley Wasserman and Katherine Faust. *Social Network Analysis. Methods and Applications*. Cambridge: Cambridge university press edition, 1994.

George Wolf and Joseph DiMona. *Frank Costello: Prime Minister of the Underworld*. Morrow, 1974.

I

NAME : Joseph BONANNO

ALIASES : Joe Bananas, Joe Bonanno, Joe Bonnano, Joe Bouventre

DESCRIPTION : Born 1-18-1905 Castellammare, Sicily, 5'9", 190 lbs, brown eyes, brown-grey hair, naturalized 5-17-45, Brooklyn, NY.

LOCALITIES FREQUENTED : Resides 1847 East Elm Street, Tucson, Arizona. Travels extensively about U.S. & makes frequent trips to Italy.

FAMILY BACKGROUND : Married Filippa LaBruzzo; daughter: Catherine; sons: Salvatore (married to Rosalie Profaci, niece of Giuseppe Profaci) and Joseph: father: Salvatore; mother: Catherine Bouventre; both parents deceased.

CRIMINAL ASSOCIATES : Lucky Luciano, Francisco Costiglia, Giuseppe Profaci, Anthony Corallo, Thomas Lucchese, Carmine Galante.

CRIMINAL HISTORY : FBI #2534540 NYCPD #B-85172 I&NS #C-6602167 Record dating from 1930 includes arrests for grand larceny, possession of gun, transportation of machine guns, obstruction of justice.

BUSINESS : Has interests in Grande Cheese Co., Fond du Lac, Wis.; Alliance Realty & Insurance, Tucson, Arizona; and Brunswick Laundry Service, Brooklyn, N.Y.

MODUS OPERANDI : Attended 1957 Apalachin Mafia meeting and Binghamton, NY, meeting 1956. One of the most important Mafia leaders in U.S. and attends all top-level Mafia meetings. Makes trips to Italy to confer with Mafia leaders there and to negotiate for international narcotic trafficking.

Figure 1: Record Number One: Joe Bonanno

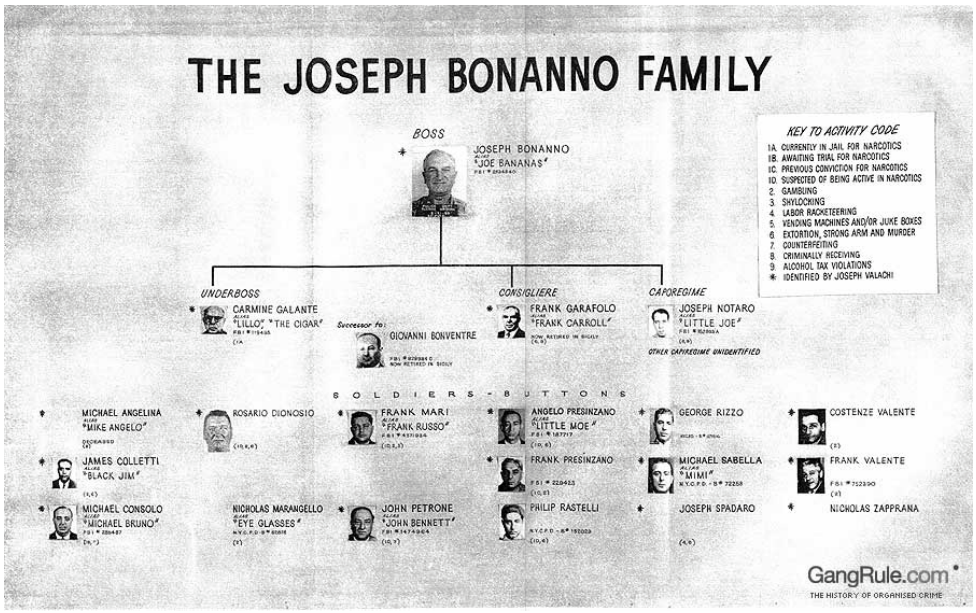

Figure 2: The Bonanno Family.

Source: Valachi's testimony in the McClellan Hearings.

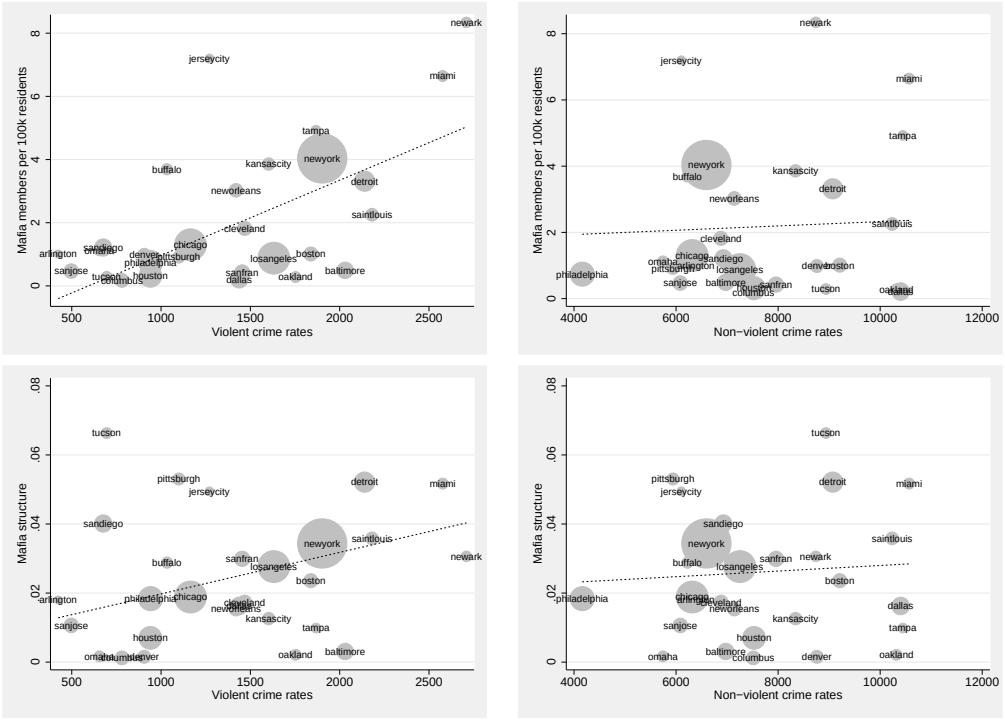

Figure 3: Violent and Non-Violent Crime Rates by Mafia Density

Notes: Mafia structure is equal to the average eigenvector index. Violent (murder, rape, robbery, and assault) and non-violent (burglary, larceny, and auto theft) crime rates are based on uniform crime reports.

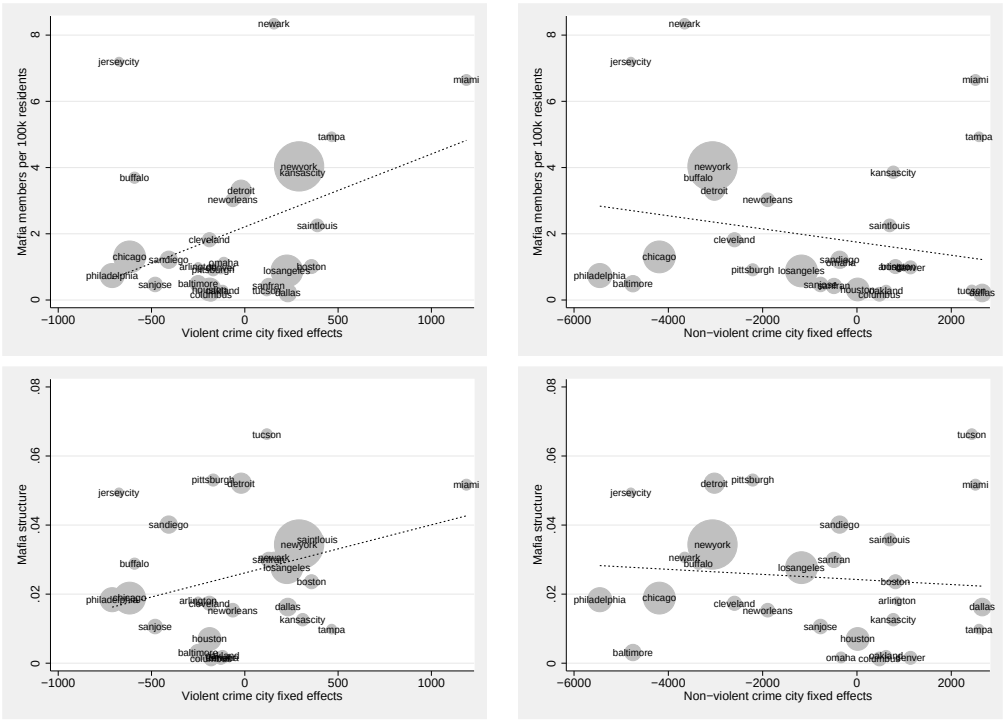

Figure 4: Violent and Non-Violent Crimes Rate Residuals by Mafia Density

Notes: Mafia structure is equal to the average eigenvector index. Violent (murder, rape, robbery, and assault) and non-violent (burglary, larceny, and auto theft) crime rates are based on uniform crime reports. The city fixed effects are based on typical crime regression that controls for several time-varying city characteristics.

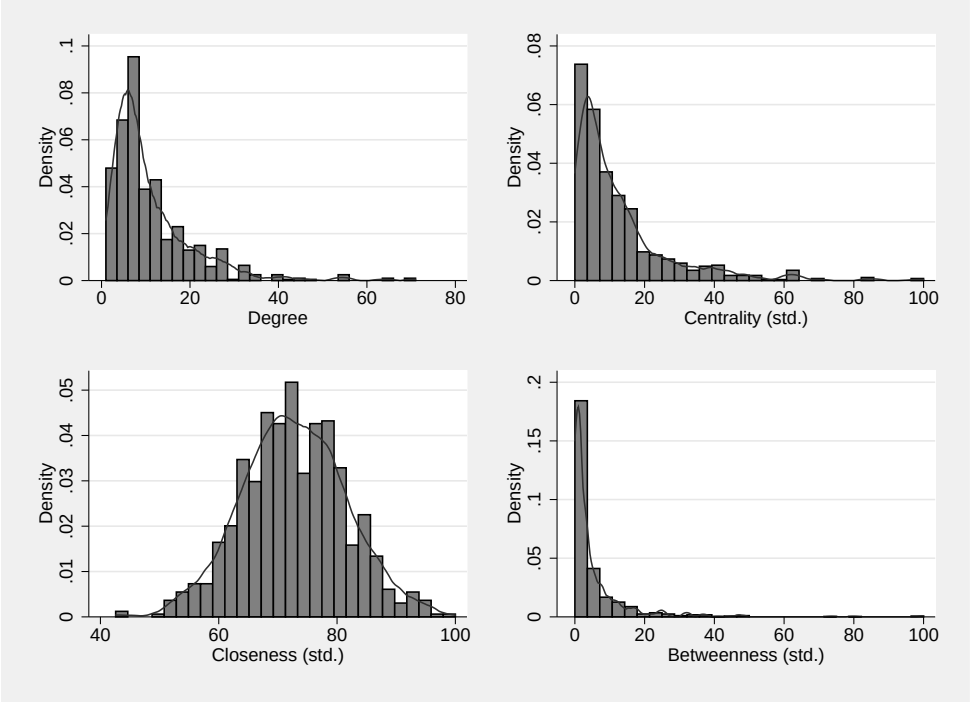

Figure 5: Density of *Degree*

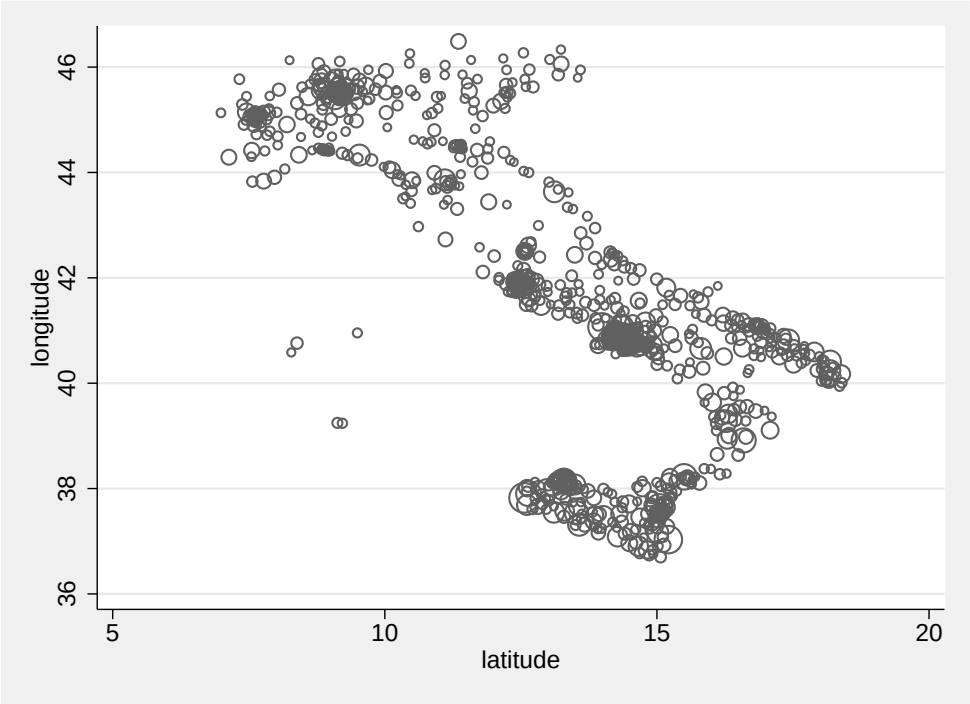

Figure 6: Geographical Distribution of Mafia Surnames.

Notes: Each circle represents a zip code. The size of the circles is proportional to the number of US Mafiamembers' surnames found in today's Italian phone directory. The plot shows only 20 percent of the distribution of surnames.

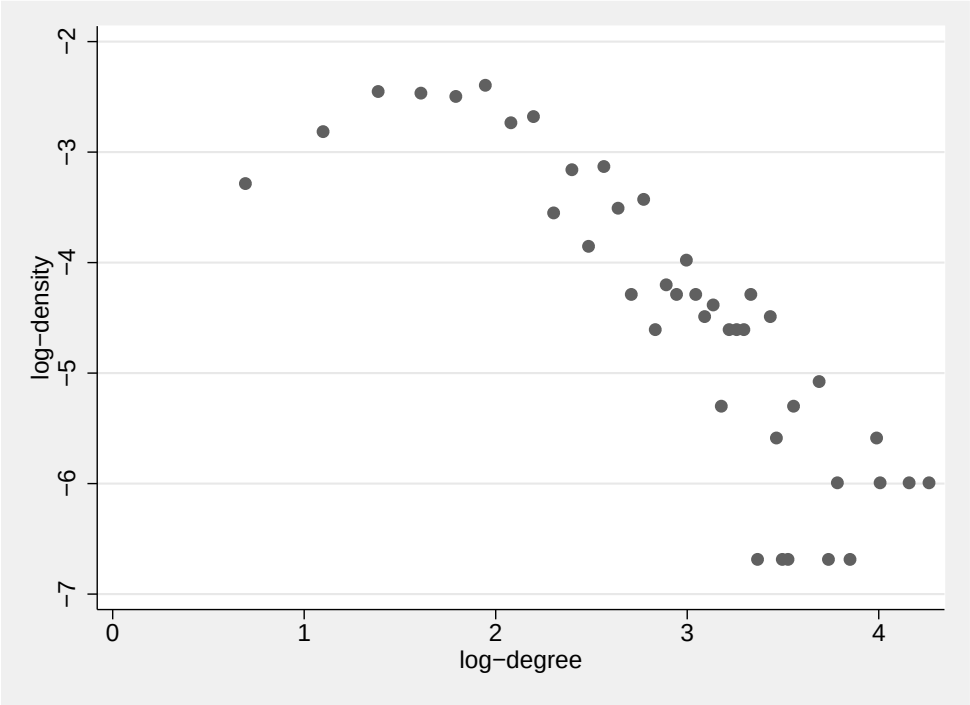

Figure 7: Log Density on log *Degree*

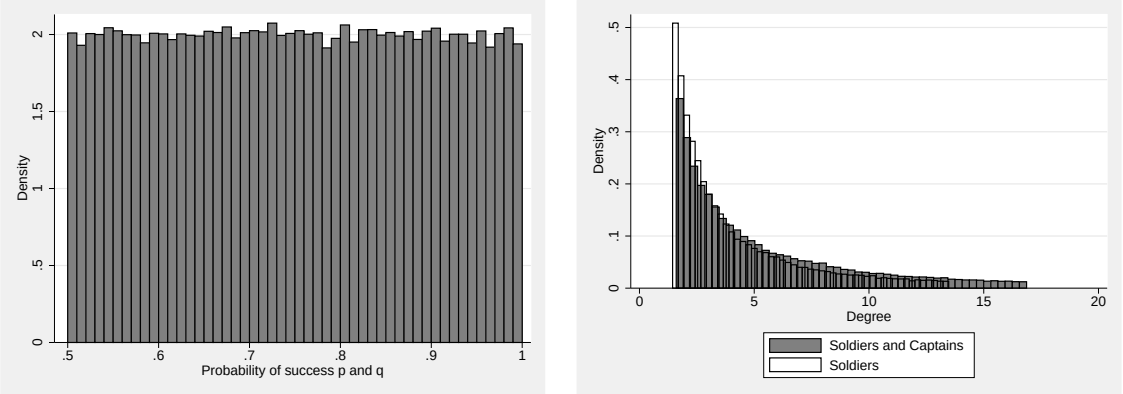

Figure 8: Simulations of *Degree*

Notes: τ is equal to 0.5, p and q are set to be the same.

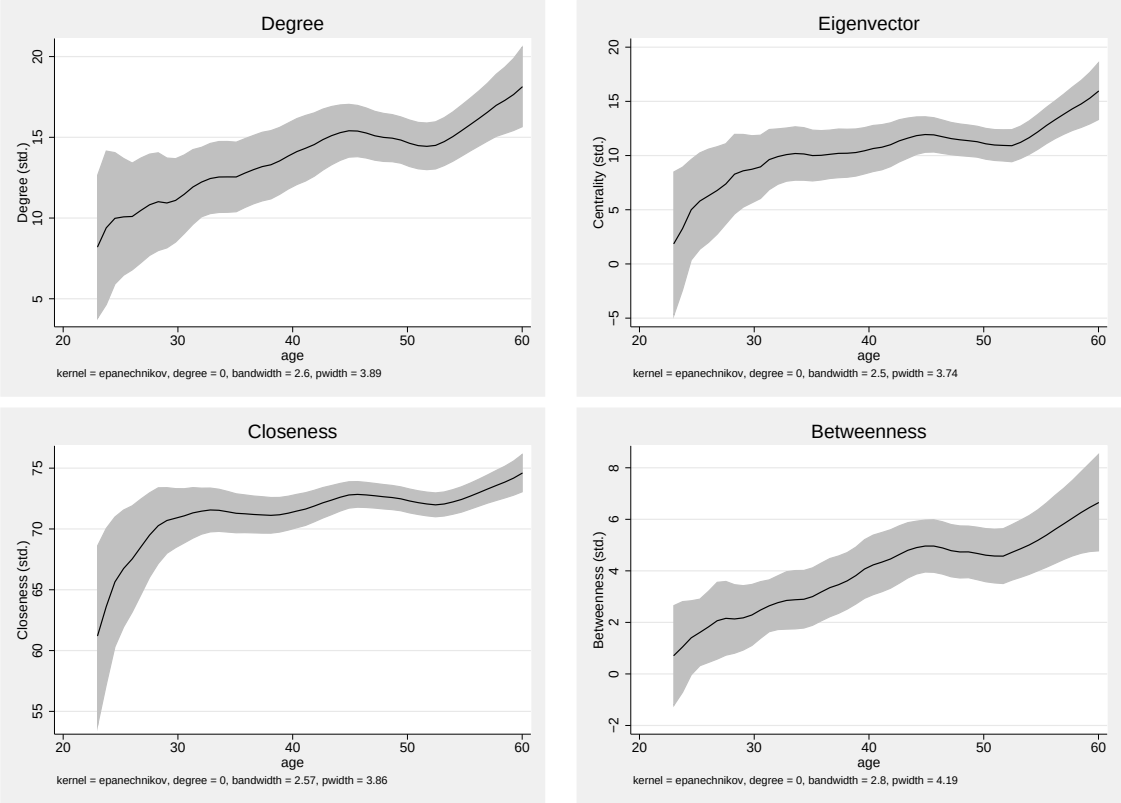

Figure 9: Centrality and Age

Notes: The figures show Kernel-weighted local polynomial regression and the corresponding 95 percent confidence interval.

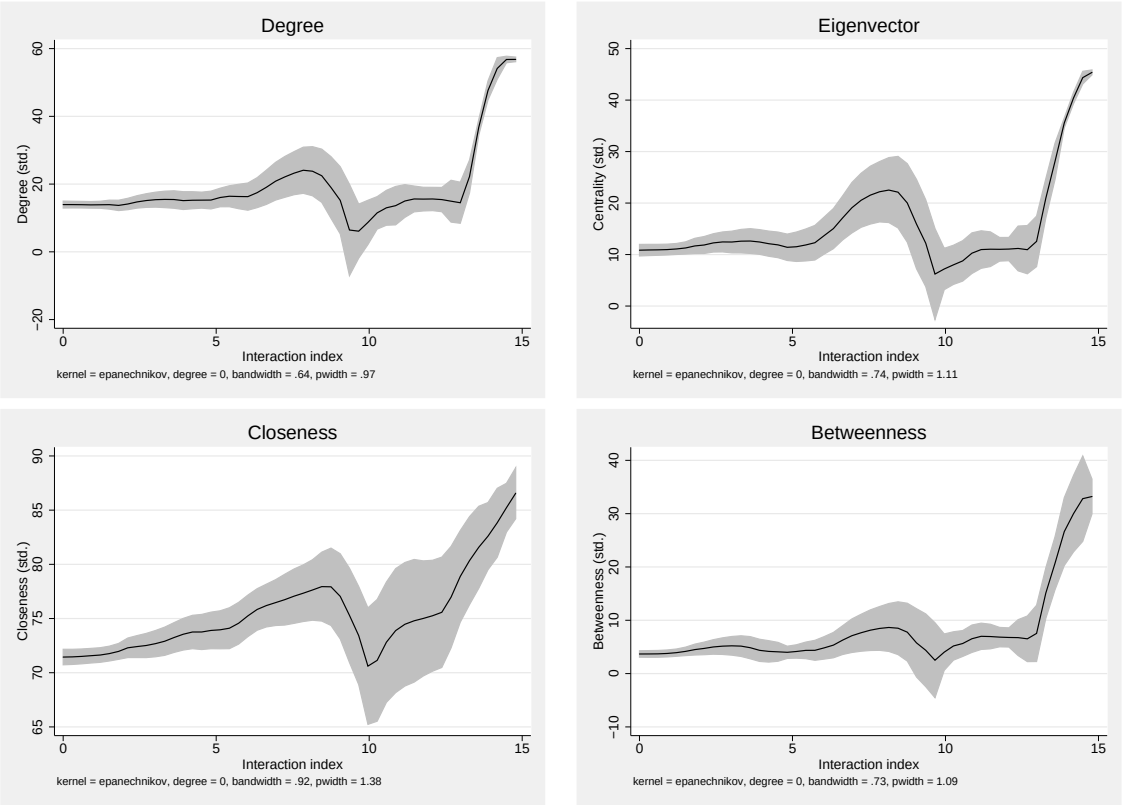

Figure 10: Centrality and "Tradition"

Notes: The figures show Kernel-weighted local polynomial regression and the corresponding 95 percent confidence interval. The interaction index has been truncated at the 95th percentile.

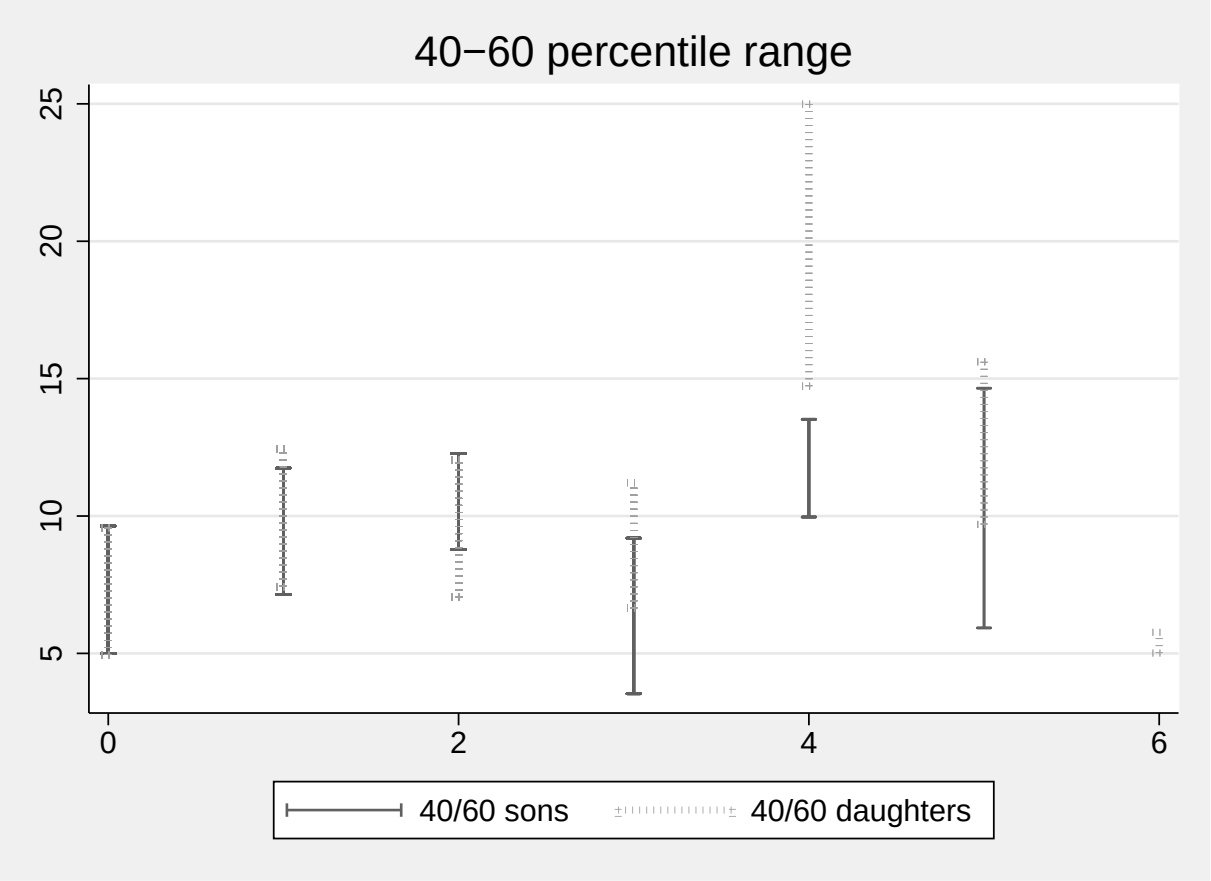

Figure 11: Centrality, Sons and Daughters

Notes: Each line represent the 40-60 percentile range.

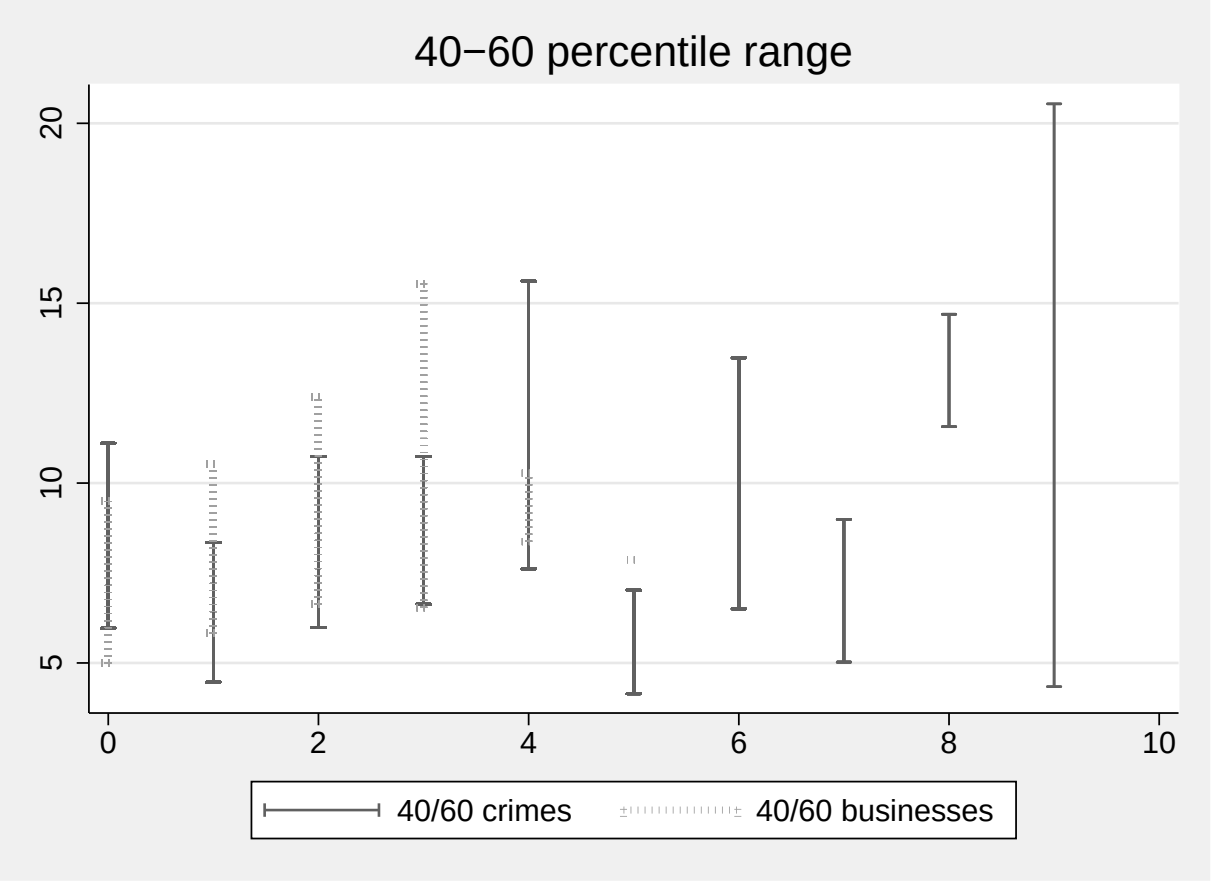

Figure 12: Centrality, Types of Crimes and of Businesses

Notes: Each line represent the 40-60 percentile range.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	N
	Cities with the Mafia		
Violent crimes	1489.415	699.841	410
Murders	24.941	12.136	410
Rapes	73.215	29.544	410
Robberies	767.827	377.486	410
Assaults	642.239	394.551	410
	Cities without the Mafia		
Violent crimes	1015.653	626.481	932
Murders	16.041	11.19	932
Rapes	66.253	31.78	932
Robberies	464.084	338.572	932
Assaults	469.275	318.857	932

Notes: Violent (murder, rape, robbery, and assault) crime rates are based on uniform crime reports between 1970 and 1992. Cities with the Mafia are cities where Families seemed to be active in 1960 according to our criminal records.

Table 2: Summary Statistics of Individual Characteristics

Variable	Mean	Std. Dev.	Min.	Max.	N
		Person			
Born in the U.S.	0.59	0.49	0	1	801
Born in Italy (except Sicily)	0.1	0.3	0	1	801
Born in Sicily	0.29	0.45	0	1	801
Year of birth	1911.48	7.66	1900	1937	623
Height in feet	5.61	0.2	5	6.25	790
Weight in pounds	176.24	26.63	95	365	789
		Family			
Married	0.75	0.43	0	1	801
Divorced	0.07	0.25	0	1	801
Connected wife	0.18	0.38	0	1	801
Number of children	1.01	1.44	0	8	801
Fraction of daughters	0.49	0.37	0	1	351
Siblings	1.96	2.1	0	11	801
Extended family members	1.62	1.04	1	6	801
		Activities			
Resides in NY	0.43	0.49	0	1	801
Resides in NJ	0.07	0.25	0	1	801
Violent crimes	0.63	0.48	0	1	801
Age at first arrest	23.19	7.59	6	57	521
Never arrested	0.16	0.37	0	1	801
Types of crime committed	2.58	1.7	0	9	801
Types of businesses	1.07	0.98	0	5	801

Table 3: Summary Statistics of Legal Businesses

Variable	Mean	Std. Dev.	Min.	Max.
Drugstores	0.27	0.46	0	2
Restaurants	0.18	0.38	0	1
Food companies	0.09	0.28	0	1
Manual laborer	0.07	0.26	0	1
Casinos	0.07	0.25	0	1
Real estate	0.05	0.23	0	1
Import export	0.05	0.22	0	1
Car dealer	0.05	0.22	0	1
N		801		

Table 4: Summary Statistics of Crimes

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
Drug offenses	0.43	0.5	0	1
Robbery	0.26	0.44	0	1
Murder	0.23	0.42	0	1
Weapon offenses	0.22	0.42	0	1
Simple assault	0.21	0.41	0	1
Larceny	0.20	0.4	0	1
Burglary	0.13	0.34	0	1
Gambling	0.13	0.33	0	1
Liquor offenses	0.13	0.34	0	1
Extortion	0.07	0.25	0	1
Counterfeiting	0.07	0.25	0	1
N		801		

Table 5: List of Criminals with the Highest and the Lowest *Degree*

Lastname	Name	*Degree*	*Top*	*High*	*Apalachin*	Rank (Family)
Criminals with the highest 15 degrees						
Lucania	Salvatore	71	1	1	0	Boss
Ormento	John	71	0	2	1	Caporegime (Lucchese)
Accardo	Settimo	64	1	2	0	Caporegime (Lucchese)
Accardo	Antonio	64	0	1	0	Boss
Genovese	Vito	55	1	0	1	Underboss (Luciano) later Boss
Genovese	Michael	55	0	1	1	Boss
Coppola	Michael	54	2	1	0	Caporegime (Genovese)
Coppola	Frank	54	0	5	0	Deported to Italy
Coppola	Stephen	54	0	1	0	Soldier (Maggadino)
Strollo	Antonio	47	2	2	0	Caporegime (Genovese)
Profaci	Frank	44	1	1	0	Boss
Profaci	Joseph	44	1	2	1	Boss (son of Frank)
Santoro	Salvatore	42	0	0	0	Underboss (Lucchese)
Vitale	Vito	40	1	1	0	Boss in Sicily
Vitale	Salvatore	40	1	0	0	Underboss (Bonanno)
Criminals with the lowest 15 degrees						
Castorina	Vincent	1	0	0	0	
Kornhauser	Max	1	0	2	0	
Bibbo	Nicholas	1	0	0	0	
Virusso	Santo	1	0	1	0	
Mandala	Nicholas	1	0	0	0	
Roberto	Dominick	1	0	1	0	
Simoni	Pierre	1	0	0	0	
Candelmo	John	1	0	1	0	
Colombo	Frank	1	0	0	0	
Bongiorno	Frank	1	0	1	0	
Amari	Philip	1	0	1	0	
Peloso	Antonio	1	0	0	0	
Labarbara	Joseph	1	0	0	0	
Pine	Grace	1	1	1	0	
Valle	Alarico	1	1	0	0	

Table 6: Summary Statistics of Network-related Variables

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
		Measures of Centrality		
Degree	11.13	9.55	1	71
Degree (std.)	15.67	13.45	1.41	100
Centrality (std.)	12.59	14.16	0.01	100
Closeness (std.)	72.74	8.76	42.57	100
Betweenness (std.)	5.09	9.47	0	100
Top ranked	0.55	0.5	0	1
Top rank citations	0.85	1.02	0	7
		Exposure to Tradition		
Interaction index	3.7	7.52	0	65.40
N		801		

Table 7: Positive Assortativity Among Crime and Business Types

Crimes			Businesses		
	$\widehat{\beta}$	SD($\widehat{\beta}$)		$\widehat{\beta}$	SD($\widehat{\beta}$)
Murder	0.440	0.083	Restaurants	0.158	0.095
Robbery	0.478	0.091	Drugstores	0.184	0.098
Simple assault	0.339	0.087	Food companies	0.447	0.111
Burglary	0.232	0.099	Real estate	0.310	0.134
Larceny	0.068	0.094	Import export	0.403	0.155
Counterfeiting	0.361	0.148	Manual laborer	0.215	0.129
Drug offenses	0.939	0.053	Casinos	0.348	0.111
Gambling	0.608	0.117			
Liquor offenses	-0.001	0.072			
Weapons offenses	0.160	0.086			
Extortion	0.403	0.135			

Notes: Estimated coefficients and clustered standard errors by surnames from a regression of crime type dummies on the fraction of associates who perpetrated the same crime type. We restrict the data to businesses held and crimes perpetrated by at least five percent of the sample.

Table 8: The Determinants of the Eigenvector Index

	(1)	(2)	(3)	(4)
	Eigenvalue index			
Extended family members	4.726***	5.539***	5.791***	5.026***
	(0.908)	(0.859)	(0.831)	(0.785)
Born in Italy (except Sicily)	-0.583			0.091
	(1.622)			(1.571)
Born in Sicily	2.282*			3.156**
	(1.284)			(1.222)
Year of birth	-0.247***			-0.228***
	(0.064)			(0.066)
Interaction index	2.409*			2.407*
	(1.409)			(1.310)
Height in feet	2.697			2.704
	(2.657)			(2.506)
Weight in pounds	0.003			-0.002
	(0.020)			(0.017)
Married		0.244		-0.691
		(1.203)		(1.160)
Divorced		0.767		-0.004
		(2.050)		(2.132)
Connected wife		5.796***		4.711***
		(1.602)		(1.518)
Number of children		1.025***		0.622*
		(0.365)		(0.349)
Fraction of daughters		0.987		1.846
		(2.023)		(1.923)
Siblings		0.051		-0.006
		(0.204)		(0.219)
Resides in NY			4.567***	4.842***
			(1.070)	(1.059)
Resides in NJ			6.200***	5.965***
			(2.118)	(2.012)
Violent crimes			1.863**	1.426
			(0.901)	(0.927)
Age at first arrest			-0.079*	-0.056
			(0.042)	(0.058)
Never arrested			-2.624*	-2.222
			(1.346)	(1.828)
Types of crime committed			0.563*	0.676**
			(0.295)	(0.318)
Types of businesses			1.981***	1.818***
			(0.541)	(0.530)
Observations	801	801	801	801
R-squared	0.200	0.204	0.226	0.287

Notes: The regressions include also missing dummies for year of birth, height, and weight. Clustered (by family) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 9: The Determinants of Importance Indices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Eigenvalue		Closeness		Betweenness	
	Degree	Total	Residual	Total	Residual	Total	Residual
Ext. family members	5.495***	5.026***	-0.121	3.704***	2.871***	0.913***	-0.009
	(0.876)	(0.785)	(0.489)	(0.416)	(0.598)	(0.319)	(0.036)
Born in Italy (ex. Sicily)	-0.576	0.091	0.630	-0.190	-0.216	0.102	0.133
	(1.463)	(1.571)	(0.771)	(1.017)	(0.849)	(0.803)	(0.134)
Born in Sicily	2.193*	3.156**	1.102*	2.678***	1.441*	1.565***	0.174*
	(1.159)	(1.222)	(0.660)	(0.751)	(0.792)	(0.537)	(0.102)
Year of birth	-0.224***	-0.228***	-0.019	-0.177***	-0.132***	-0.063**	-0.018***
	(0.062)	(0.066)	(0.036)	(0.039)	(0.048)	(0.030)	(0.005)
Interaction index	2.751**	2.407*	-0.170	1.431***	3.495***	0.034	-0.055
	(1.270)	(1.310)	(0.806)	(0.512)	(0.847)	(0.473)	(0.055)
Height in feet	3.648	2.704	-0.714	3.316**	1.559	1.463	-0.037
	(2.368)	(2.506)	(1.258)	(1.563)	(1.678)	(1.063)	(0.205)
Weight in pounds	0.004	-0.002	-0.006	-0.001	0.003	-0.003	0.001
	(0.018)	(0.017)	(0.008)	(0.012)	(0.013)	(0.009)	(0.001)
Married	-0.122	-0.691	-0.577	0.339	-0.538	0.401	0.049
	(1.102)	(1.160)	(0.557)	(0.694)	(0.810)	(0.509)	(0.094)
Divorced	-0.247	-0.004	0.227	0.094	0.775	0.220	0.124
	(2.000)	(2.132)	(0.978)	(1.408)	(1.441)	(0.959)	(0.151)
Connected wife	4.422***	4.711***	0.569	1.858**	1.845**	-0.388	0.092
	(1.391)	(1.518)	(0.737)	(0.744)	(0.924)	(0.537)	(0.104)
Number of children	0.580*	0.622*	0.078	0.319*	0.389	0.025	-0.001
	(0.332)	(0.349)	(0.174)	(0.190)	(0.263)	(0.137)	(0.028)
Fraction of daughters	0.734	1.846	1.159	0.755	-0.084	0.382	0.169
	(1.580)	(1.923)	(0.984)	(0.997)	(1.055)	(0.723)	(0.139)
Siblings	0.181	-0.006	-0.176	0.175	0.250*	0.083	0.038**
	(0.195)	(0.219)	(0.112)	(0.126)	(0.133)	(0.090)	(0.020)
Resides in NY	2.284**	4.842***	2.703***	4.336***	0.239	3.176***	0.195**
	(0.892)	(1.059)	(0.628)	(0.596)	(0.575)	(0.418)	(0.084)
Resides in NJ	3.139	5.965***	3.025***	5.783***	2.644	4.189***	0.364**
	(2.087)	(2.012)	(0.881)	(1.090)	(1.735)	(0.819)	(0.179)
Violent crimes	0.577	1.426	0.886*	0.268	0.426	-0.025	0.065
	(0.854)	(0.927)	(0.514)	(0.541)	(0.621)	(0.410)	(0.081)
Age at first arrest	-0.045	-0.056	-0.014	-0.065	-0.030	-0.042	-0.011**
	(0.051)	(0.058)	(0.035)	(0.041)	(0.029)	(0.031)	(0.005)
Never arrested	-2.239	-2.222	-0.124	-1.971*	-1.322	-0.834	-0.219
	(1.628)	(1.828)	(1.067)	(1.187)	(1.033)	(0.862)	(0.150)
Types of crime committed	0.572**	0.676**	0.140	0.187	0.206	-0.103	-0.004
	(0.288)	(0.318)	(0.158)	(0.209)	(0.177)	(0.137)	(0.026)
Types of businesses	1.346***	1.818***	0.558**	1.051***	0.765**	0.368*	0.088**
	(0.479)	(0.530)	(0.232)	(0.292)	(0.332)	(0.201)	(0.040)
Observations	801	801	801	801	801	801	801
R-squared	0.322	0.287	0.069	0.370	0.266	0.134	0.082

Notes: The regressions include also missing dummies for year of birth, height, and weight. The leader variable counts the number of words that describe mobsters as leaders. Clustered (by family) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 10: Robustness Checks

	(1)	(2)	(3)	(4)
			Eigenvalue	
	Eigenvalue	Leadership	Above median	Below median
Extended family members	5.026***	-0.009	3.147***	0.447
	(0.785)	(0.036)	(0.845)	(0.426)
Born in Italy (except Sicily)	0.091	0.133	-2.989	0.260
	(1.571)	(0.134)	(3.038)	(0.418)
Born in Sicily	3.156**	0.174*	-0.229	-0.099
	(1.222)	(0.102)	(1.858)	(0.312)
Year of birth	-0.228***	-0.018***	-0.163	-0.021
	(0.066)	(0.005)	(0.115)	(0.016)
Interaction index	2.407*	-0.055	2.203	0.630
	(1.310)	(0.055)	(1.408)	(0.451)
Height in feet	2.704	-0.037	0.123	0.797
	(2.506)	(0.205)	(3.969)	(0.601)
Weight in pounds	-0.002	0.001	-0.015	0.005
	(0.017)	(0.001)	(0.038)	(0.005)
Married	-0.691	0.049	-2.699	0.184
	(1.160)	(0.094)	(2.284)	(0.261)
Divorced	-0.004	0.124	-0.537	-0.333
	(2.132)	(0.151)	(3.927)	(0.484)
Connected wife	4.711***	0.092	6.958***	0.443
	(1.518)	(0.104)	(2.195)	(0.313)
Number of children	0.622*	-0.001	0.633	0.014
	(0.349)	(0.028)	(0.509)	(0.079)
Fraction of daughters	1.846	0.169	2.108	0.370
	(1.923)	(0.139)	(3.129)	(0.414)
Siblings	-0.006	0.038**	-0.045	0.021
	(0.219)	(0.020)	(0.325)	(0.059)
Resides in NY	4.842***	0.195**	3.672**	0.561**
	(1.059)	(0.084)	(1.788)	(0.268)
Resides in NJ	5.965***	0.364**	3.716	1.852***
	(2.012)	(0.179)	(2.970)	(0.494)
Violent crimes	1.426	0.065	3.524**	-0.025
	(0.927)	(0.081)	(1.568)	(0.223)
Age at first arrest	-0.056	-0.011**	0.020	-0.020
	(0.058)	(0.005)	(0.097)	(0.016)
Never arrested	-2.222	-0.219	-2.979	-0.960**
	(1.828)	(0.150)	(2.882)	(0.448)
Types of crime committed	0.676**	-0.004	1.058**	-0.032
	(0.318)	(0.026)	(0.484)	(0.078)
Types of businesses	1.818***	0.088**	2.256***	-0.009
	(0.530)	(0.040)	(0.858)	(0.114)
Observations	801	801	401	400
R-squared	0.287	0.082	0.195	0.114

Notes: The dependent variable "Leadership" counts the number of times the words "boss," "highest," "most," "head," 'top," "high," "influential," "important," "leader," "leading," "powerful," and "representing" are cited. The regressions include also missing dummies for year of birth, height, and weight. Clustered (by family) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Regressions of *Indegree* with Random Truncation of Half of the Sample

	β	se(β)	simulated β	
			average	median
Extended family members	5.47	0.87	5.97	5.83
Born in Italy (except Sicily)	-0.49	1.45	-0.61	-0.66
Born in Sicily	2.15	1.15	2.14	2.24
Year of birth	-0.22	0.06	-0.23	-0.23
Interaction index	2.80	1.26	3.23	2.93
Height in feet	3.66	2.36	3.60	3.89
Weight in pounds	0.00	0.02	0.00	0.00
Married	-0.10	1.10	-0.05	-0.13
Divorced	-0.22	1.99	-0.21	-0.25
Connected wife	4.28	1.37	4.59	4.43
Number of children	0.57	0.33	0.60	0.59
Fraction of daughters	0.64	1.57	0.72	0.79
Siblings	0.19	0.19	0.18	0.19
Resides in NY	2.24	0.88	2.47	2.40
Resides in NJ	3.02	2.06	3.15	3.08
Violent crimes	0.54	0.84	0.55	0.48
Age at first arrest	-0.04	0.05	-0.05	-0.05
Never arrested	-2.18	1.63	-2.23	-2.12
Types of crimes committed	0.56	0.29	0.57	0.57
Types of businesses	1.30	0.47	1.36	1.38

Notes: The first two columns show the coefficients and the standard errors based on the whole sample. Columns 3 and 4 show the mean and the median coefficients of 500 regressions based on randomly truncated samples.