

Review of Network Economics

Volume 11, Issue 3

2012

Article 10

Organized Crime Networks: an Application of Network Analysis Techniques to the American Mafia

Giovanni Mastrobuoni, *Collegio Carlo Alberto*
Eleonora Patacchini, *La Sapienza Università di Roma,*
EIEF, CEPR and IZA

Recommended Citation:

Mastrobuoni, Giovanni and Patacchini, Eleonora (2012) "Organized Crime Networks: an Application of Network Analysis Techniques to the American Mafia," *Review of Network Economics*: Vol. 11: Iss. 3, Article 10.
DOI: 10.1515/1446-9022.1324

©2012 De Gruyter. All rights reserved.

Organized Crime Networks: an Application of Network Analysis Techniques to the American Mafia

Giovanni Mastrobuoni and Eleonora Patacchini

Abstract

Using a unique data set on criminal profiles of 800 US Mafia members active in the 1950s and 1960s and on their connections within the Cosa Nostra network, we use simple network analysis techniques to document the structure and composition of the geometry of criminal ties between mobsters. The use of different network centrality measures allows us to collect evidence in line with so far only conjectured views on the functioning of the Mafia. In particular, we shed light on the extent to which family relationships, community roots and ties, legal and illegal activities predict the criminal ranking of the “men of honor,” suggesting the main characteristics that can be used to detect criminal leaders. Our results are remarkably in line with the evidence that mafia organizations tend to be extremely hierarchical.

KEYWORDS: network centrality indices, intermarriage, assortative matching, crime

Author Notes: We are thankful to Theo Diasakos, Matthew Jackson, Claudio Lucifora, Franco Peracchi, Jesse Rothstein, Rocco Sciarrone, Serena Uccello, Aleksey Tetenov and seminar participants at the Economics seminar at University of Arizona, the Labor and Development seminar at Princeton University, the Public Policy seminar at Berkeley, and the workshop on the Economics of Crime and Organized Crime in Palermo, on Institutions, Individual Behavior and Economic Outcomes in Alghero, and on the one in Applied Economics 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 is grateful to the Italian Academy at Columbia University for their hospitality.

1 Introduction

Despite more than 40 years of law enforcement's success in the fight against organized crime, the Mafia has continued following many of the same rules, and is still active in many countries, including the US. According to the FBI¹, in 2005 there were 651 pending investigations related to the Italian-American Mafia; 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 no longer holds full control of racketeering. With the end of the Cold War and the advent of globalization, "transnational" organized crime organizations 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 gross domestic product (Schneider and Enste, 2000, Wagley, 2006). Williams (2001) discusses how networks within and across these organizations facilitate their fortunes.

Notwithstanding the magnitude of these numbers, the illicit nature of organized crime activities has precluded empirical analysis and the literature has overwhelmingly been anecdotal or theoretical (Reuter, 1994, Williams, 2001).²

In this paper, we exploit individual level data on 800 Mafia members that were active just before the 1961 crackdown to study the factors underlying criminal connections and hierarchies, ranking mobsters based on the number and the quality of their connections. The data source are the Federal Bureau of Narcotics (FBN) files on American Mafia members that were active and alive in 1960 (MAF, 2007).³

Recent papers on social networks show that an individual's position within

¹The source is www.fbi.gov.

²Levitt and Venkatesh (2000) use detailed financial activities of a drug-selling street gang to analyze gang behavior. But most gangs do not appear to engage in crimes motivated and organized according to formal-rational criteria (Decker et al., 1998).

³At the end of the 1950s the FBN, which 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).

a network is indeed crucial in explaining her or his level of activity (Ballester et al., 2006). If individual decisions are related 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 relationships, evidence documenting patterns of association between agents is a first priority, partly because it can inform the theoretical literature looking at network formation processes. Several models of network formation have recently been proposed. Most rely on some form of pairwise regressions (see, for example, Bramoullé and Fortin, 2010). Given the hierarchical nature of the network at hand (the Mafia), in this study we focus on the factors that are related to the centrality of mobsters rather than on those determining single connections.

Empirical evidence along this line is helpful to understand organizational rules of the Mafia, as well as to identify the “key players.” We seek to answer questions such as: to what extent 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 are associated with the most central agents in the network? The U.S. experience with Cosa Nostra and the wealth of individual-level information of our data may offer clues to promising control techniques in countries where organized crime is on the rise (Jacobs and Gouldin, 1999).

The Mafia network is probably the best example of real-world network where the geometry of connections are crucial for understanding the activity of the network as these connections are the building blocks of the entire Mafia, and more generally of organized crime even today. Valachi’s 1963 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).⁴ Connections are thus

⁴The remaining 9 rules are: never look at the wives of friends, never be seen with cops, don’t go to pubs and clubs, always being available for Cosa Nostra is a duty - even if one’s wife is going through labor, appointments must be strictly respected, wives must be treated with respect, only truthful answers must be given when asked for information by another member, money cannot be appropriated if it belongs to others or to other families, certain types of people can’t be part of Cosa Nostra (including anyone who has a close relative in

necessary for a criminal career within the Mafia. Moreover, leadership positions aren't simply inherited; soldiers elect their bosses using secret ballots (Falcone and Padovani, 1991, pg. 101).

Francisco Costiglia, alias Frank Costello, a Mafia boss 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 this research is to shed light on 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, this is the first study to use 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.

Several measures on how central and important members are within the network are employed. A mobster might exert power because he is either connected to a large number of members or connected to a few 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. One of the advantages of indices that measure importance based on network-based connections over indices that rely on rank in the organizational hierarchy is that they measure importance in a continuous manner—Mafia bosses are not all equally powerful, and *capiregime* often differ in their level of importance—, and are more robust to classification errors (Klerks, 2003).

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 are more likely to be associated with criminals who operate similar illegal businesses if they try to build cartels. Or they may try to diversify the risk of detection by 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. These marriages may thus

the police, anyone with a two-timing relative in the family, anyone who behaves badly and does not possess moral values).

have the same function as alliances among European royalty: providing new protectors. Children may also prove important, both because of strategic endogamy and because male descendants represent trusted potential associates.⁵ Another way children or, more generally, relatives might increase the trust towards 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 a Mafia clan's 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 their monopoly power in exchange for part of their revenues. Large clans will therefore be more powerful but also more exposed.

Some of these 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 a 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) look at the optimal internal structure (cells versus hierarchies). Sparrow (1991) and later Coles (2001) propose the use of network analysis to study criminal networks. Morselli (2003) follows their proposal analyzing connections within a single New York based family (the Gambino family), Krebs (2002) analyzes connections among the September 2001 hijackers' terrorist cells, Baker and Faulkner (1993) study the social organization of three well-known price-fixing conspiracies in the heavy electrical equipment industry, Natarajan (2000, 2006) analyzes wiretap conversations among drug dealers, McGloin (2005) analyzes the connections among gang members in Newark (NJ), Sarnecki (1990, 2001) uses network analysis to study co-offending behavior among Swedish teenagers. These studies highlight the importance of deep ties, but have little or no background information on the

⁵Another 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).

individuals.

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 the same applies to criminal behavior. Bertrand and Schoar (2006) present a macro-type analysis about the importance of family values for economic growth, and conclude with the comment that more research is needed to understand how family values shape the organization of businesses and their efficiency.

2 The Invisible Hand of the Mafia

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 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).⁶ Moreover, crimes like racketeering, drug trafficking, gambling, bootlegging, etc require a group to be properly “organized.”⁷

Empirical evidence on the relationship between crime and typically unobserved criminal organizations is, however, scarce, to say the least.⁸ 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)

⁶In his very influential theory of differential associations, (Sutherland, 1947) emphasizes that the role of gangs is to facilitate the learning of crime technology.

⁷See Alexander (1997) for an economic analysis of racketeering with collusion.

⁸Among the few exceptions, Bandiera (2003) analyzes the origins of the Sicilian Mafia, highlighting how land fragmentation, absence of rule of law, and predatory behavior generated a demand for private protection.

crimes. Crime rates are based on city-level uniform crime reports.⁹

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).

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.

Among the Mafia cities the top panels of Figure 1 shows the raw correlation between the number of Mafia members per 100,000 inhabitants and local levels

⁹We 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>

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.¹⁰ It appears 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 2 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 (city fixed effects, for example, cannot be fixed over time), 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. In the next sections of this paper, we present an analysis of the network structure within *Cosa Nostra* that can help prosecutors to understand how *Cosa Nostra* works and how to best fight it.

¹⁰This index will be used in our analysis (Sections 4 and 5). 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 A we discuss this index in more detail. Google's page rank is a variant of the eigenvector centrality index (Bonacich, 1987).

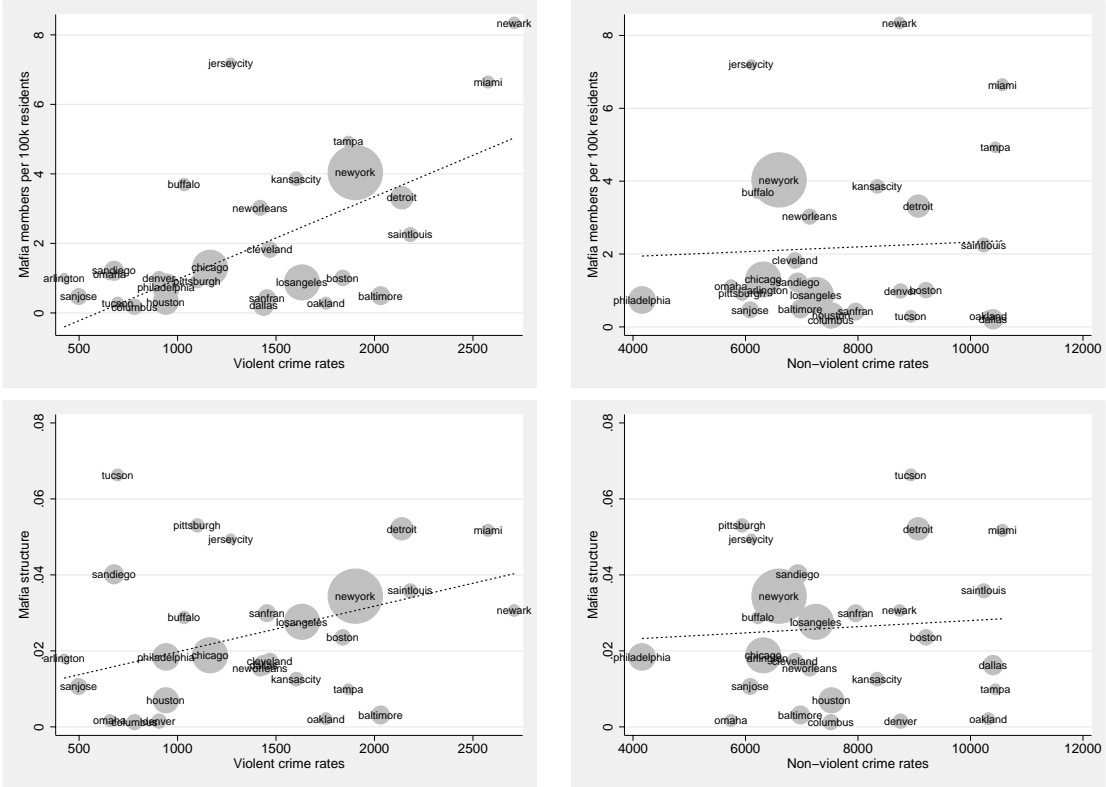


Figure 1: Violent and Non-Violent Crime Rates by Mafia Density

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

Mastrobuoni and Patacchini: Organized Crime Networks

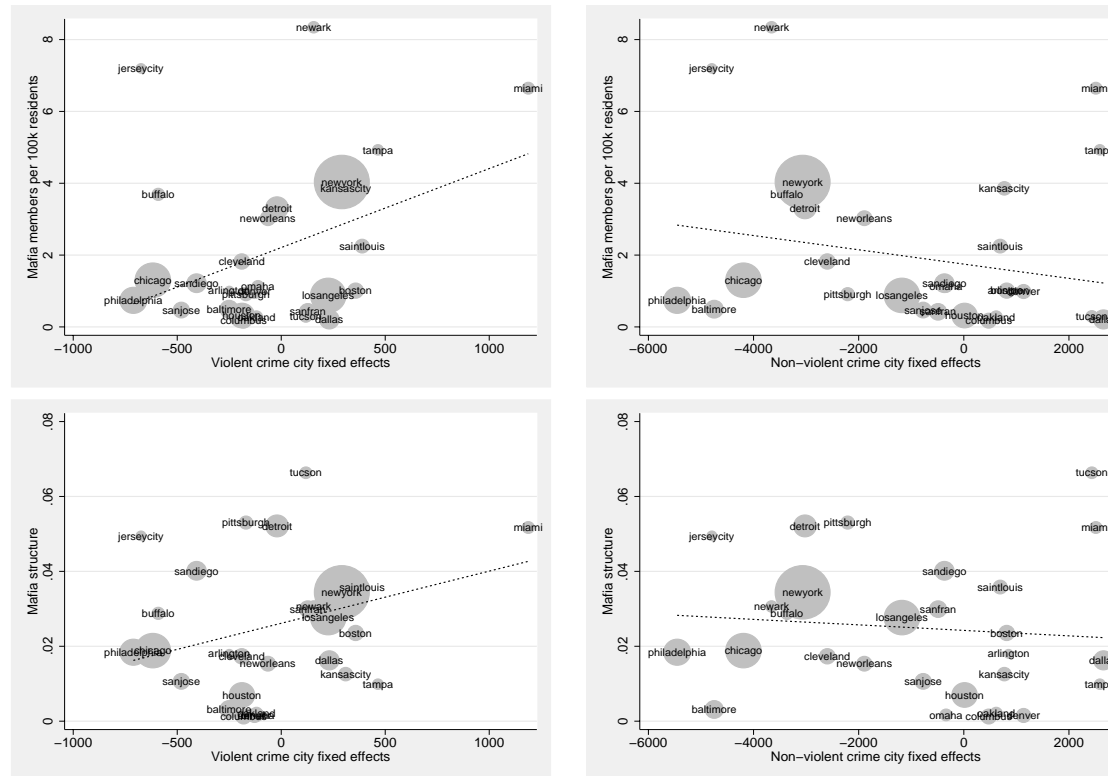


Figure 2: Violent and Non-Violent Crimes Rate Residuals by Mafia Density

Notes: Mafia structure is measured by 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.

3 Data and Descriptive Evidence

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 contain detailed information on about 800 mafia members (McWilliams, 1990). Given that in the U.S. there were an estimated 5,000 members active during those years the list represents a clearly 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. We will address the issue about the non-random nature of our sample in Section 6.

Let us begin our investigation by describing the information contained in the criminal records on mafia members. 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 refers to what the authorities knew in 1960.¹¹ The records do not report any deaths, 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 family.¹² Table 2 shows that the average age is 52 years, with the oldest mobster being 81 years old, and the youngest one 23. 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 this was a prerequisite to become a member. 73 percent of members are married, but only 60 percent of these are reported to have children. 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 shares her maiden name with some other member

¹¹Additional 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."

¹²His brother Anthony "Tough Tony," instead, was killed in 1963 and is in the records.

Table 2: Summary Statistics of Individual Characteristics

	Mean	Std. Dev.	Min	Max	Obs
Variables related to the PERSON					
Born in the U.S.	0.59	0.49	0	1	801
Born in Italy (except Sicily)	0.19	0.39	0	1	801
Born in Sicily	0.29	0.45	0	1	801
Age	52.17	10.04	23	81	801
Height in feet	5.61	0.20	5	6.25	790
Weight in pounds	176	27	95	365	790
Variables related to the FAMILY					
Interaction index	2.67	5.83	0	53.91	801
Married	0.73	0.44	0	1	801
Divorced	0.05	0.22	0	1	801
Connected wife	0.18	0.38	0	1	801
Number of children	1.02	1.44	0	8	801
Fraction of daughters	0.49	0.37	0	1	352
Siblings	1.97	2.11	0	11	801
Extended family members	1.62	1.05	1	6	801
Variables related to the ACTIVITIES					
Resides in NY	0.43	0.50	0	1	801
Resides in NJ	0.06	0.24	0	1	801
Violent crimes	0.63	0.48	0	1	801
Age of first arrest	25.02	9.06	8	67	688
Never arrested	0.16	0.37	0	1	801
Types of crimes committed	2.58	1.70	0	9	801
Types of businesses	1.07	0.97	0	5	801

(*Connected wife*). These marriages are presumably endogamous within the Mafia.¹³ The FBN reports an average of 1.97 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.¹⁴ Mobsters' criminal career starts early. They are on average 25 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. Although we do not know the total number of crimes committed by the mobsters, 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 the number of different legal businesses they have interest in. This number varies between 0 and 5 and is on average equal to 1.

Another variable summarized in Table 2 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. Looking at Figure 3 helps explaining the index. It shows the current distribution at the zip code level of the members' surnames in Italy's phone directory.¹⁵ 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 migratory 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:

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

¹³Observe that we are probably understating the percentage of marriages within the Mafia as some Mafia surnames might be missing in the data. 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.

¹⁴As a note, 18 percent of the mobsters are obese and 58 percent overweight.

¹⁵Unfortunately we could not find the distribution of surnames in 1960.

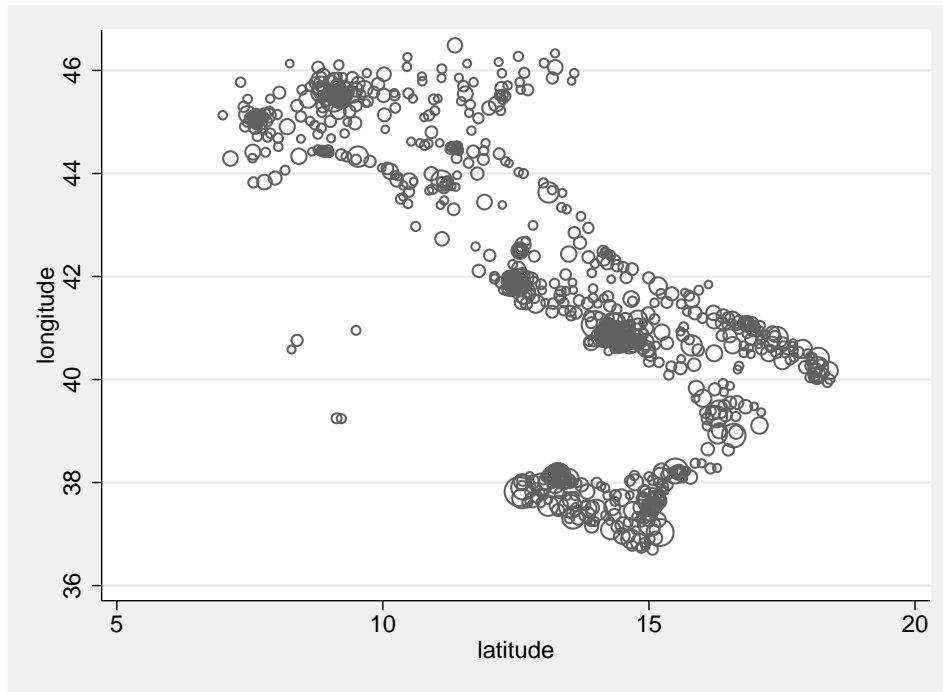


Figure 3: Geographical Distribution of Mafia Surnames.

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.

The advantage of this index is that we can compute 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 2.67 per 100,000, though it's lower when weighting the data (2.01). 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. Such an index will be used in our analysis to measure the importance of community roots and ties, the so-called "mafia culture."

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, drug-stores or were otherwise involved with the supply of food. Real estate, casinos, car dealerships, and import-export were also common businesses. Drug

Table 3: Summary Statistics of Legal Businesses

	Mean	Std. Dev.
Drugstores	0.18	0.38
Restaurants	0.09	0.31
Food companies	0.09	0.28
Manual laborer	0.07	0.26
Casinos	0.07	0.25
Real estate	0.05	0.23
Import export	0.05	0.22
Car dealer	0.05	0.22

Table 4: Summary Statistics of Crimes

	Mean	Std. Dev.
Drug offenses	0.43	0.50
Robbery	0.26	0.44
Murder	0.23	0.42
Weapon offenses	0.22	0.42
Simple assault	0.21	0.41
Larceny	0.20	0.40
Burglary	0.13	0.34
Gambling	0.13	0.33
Liquor offenses	0.13	0.34
Extortion	0.07	0.25
Counterfeiting	0.07	0.25

trafficking is the most common of the illegal activities (43 percent), perhaps partly 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. Some crimes that are typically associated with organized crime, like gambling, extortions, and liquor offences (during prohibition) are highly represented as well.


4 Network-based Measures of Importance

For our purposes, the most valuable information of these data is that each criminal record contains a list of criminal associates.

We define two mobsters to be connected whenever at least one mobster lists the other mobster's last name in his record. In other words, we construct a (symmetric) undirected network of connections, assuming that if one mobster nominate (or is nominated by) another mobster interactions are at work even though the nominations are not reciprocal. Indeed, the list of associates is restricted to at most eight names.

Mastrobuoni and Patacchini: Organized Crime Networks

There is no evidence about how the FBN established such associations, but it seems apparent that each profile lists its highest associates. Notice that the mere keeping track of connections shows that even the FBN understood how important these connections were.¹⁶ Figure 4 indicates, for example, that Joe Bonanno was associated with Luciano, Costello, Profaci, Corallo, Lucchese, and Galante. Indirect connections are clearly more numerous, as mobsters can be listed as associates in several records.

<p>NAME : Joseph BONANNO</p> <p>ALIASES : Joe Bonanno, Joe Bonanno, Joe Bonanno, Joe Bouventre</p> <p>DESCRIPTION : Born 1-18-1905 Castellana, Sicily, 5'9", 190 lbs, brown eyes, brown-grey hair, naturalized 5-17-45, Brooklyn, NY.</p> <p>LOCALITIES FREQUENTED : Resides 1847 East Elm Street, Tucson, Arizona. Travels extensively about U.S. & makes frequent trips to Italy.</p> <p>FAMILY BACKGROUND : Married Filippa LaRusso; daughter: Catherine; sons: Salvatore (married to Rosalie Profaci, niece of Giuseppe Profaci) and Joseph; father: Salvatore; mother: Catherine Bouventre; both parents deceased.</p> <p>CRIMINAL ASSOCIATES : Lucky Luciano, Francisco Costiglia, Giuseppe Profaci, Anthony Corallo, Thomas Lucchese, Carmine Galante.</p> <p>CRIMINAL HISTORY : FBI #2534540 NYCD #B-85172 IRMS #C-6602167 Record dating from 1930 includes arrests for grand larceny, possession of gun, transportation of machine guns, obstruction of justice.</p> <p>BUSINESS : Has interests in Grande Cheese Co., Fond du Lac, Wis.; Alliance Realty & Insurance, Tucson, Arizona; and Brunswick Laundry Service, Brooklyn, N.Y.</p> <p>MOBIS 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.</p>	
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------

Property of U. S. Govt - For Official Use Only
May not be disseminated or contents disclosed
Without permission of Commissioner of Narcotics

Figure 4: Record Number One: Joe Bonanno

The number of direct connections, called *degree* in network analysis, is the easiest way to measure the importance of members.

Evidence on its validity as an indicator of importance within the mafia organization can be found in Table 5.

¹⁶We will address issues related to the direct nature of our data and to the bounded

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

Lastname	Name	<i>Degree</i>	<i>Top</i>	<i>High</i>	<i>Apalachin</i>
Criminals with the highest 15 degrees					
Lucania	Salvatore	71	1	1	0
Ormento	John	71	0	2	1
Accardo	Settimo	64	1	2	0
Accardo	Antonio	64	0	1	0
Genovese	Vito	55	1	0	1
Genovese	Michael	55	0	1	1
Coppola	Michael	54	2	1	0
Coppola	Frank	54	0	5	0
Coppola	Stephen	54	0	1	0
Strollo	Antonio	47	2	2	0
Profaci	Frank	44	1	1	0
Profaci	Joseph	44	1	2	1
Santoro	Salvatore	42	0	0	0
Vitale	Vito	40	1	1	0
Vitale	Salvatore	40	1	0	0
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 5 shows the list of members with the 10 highest and 10 lowest number of direct connections.¹⁷ For each member, the table also reports qualitative information on their importance within the organization. Indeed, for each mobster the FBN records contain a paragraph about his activities within *Cosa Nostra*. 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 Upstate New York. Table 5 reveals that 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.

Although the number of direct connections (*degree*) is a simple (and valid) indicator of the importance of network members, in recent years social network studies have proposed different centrality measures to account for the variability in network location across agents, and there is no systematic criterium to pick up the “right” centrality measure for each particular situation (Borgatti, 2003, Wasserman and Faust, 1994).¹⁸

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.¹⁹ 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 measures the number of times a node

nature of the nomination number in Section 6.

¹⁷Since 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.

¹⁸See also Sparrow (1991) for a discussion on centrality indices in criminal networks.

¹⁹As first 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.

is on the shortest path between two randomly chosen nodes, and is a good measure for the member’s capacity to act like a bridge between clusters of members. A detailed exposition of these indicators can be found in Appendix A.

These measures of individual centrality allow us to consider different nuances in the definition of a Mafia leader. We also construct a qualitative indicator of importance equal to the sum of the citations *Top* and *High* discussed above (*Top Ranked Citations*), and another one which is simply a dummy equal to one when *Top Ranked Citations* is positive, called *Top Ranked*.

Table 6 collects descriptive statistics about various indicators of importance.²⁰ Fifty-five percent of the mobsters appear to be high-ranked (Table 6), meaning that their descriptions includes at least a word that implies leadership.

Table 6: Summary Statistics of Network-related Variables

Variable	Mean	Std. Dev.
Degree	11.14	9.55
Degree (std.)	14.49	13.64
Centrality (std.)	12.57	14.16
Closeness (std.)	52.53	15.25
Betweenness (std.)	5.10	9.46
Top ranked	0.55	0.50
Top rank citations	0.85	1.02

Figure 5 illustrate the empirical distribution of the network indices in our sample. Figure 5 shows that the density of *degree* is positively skewed. The eigenvector index (centrality) has a density that is very similar to that of *degree*, while the density 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 represent bridges between subsets of the network, most likely Families. Different measures of centrality thus seem to capture different roles within the criminal organization.

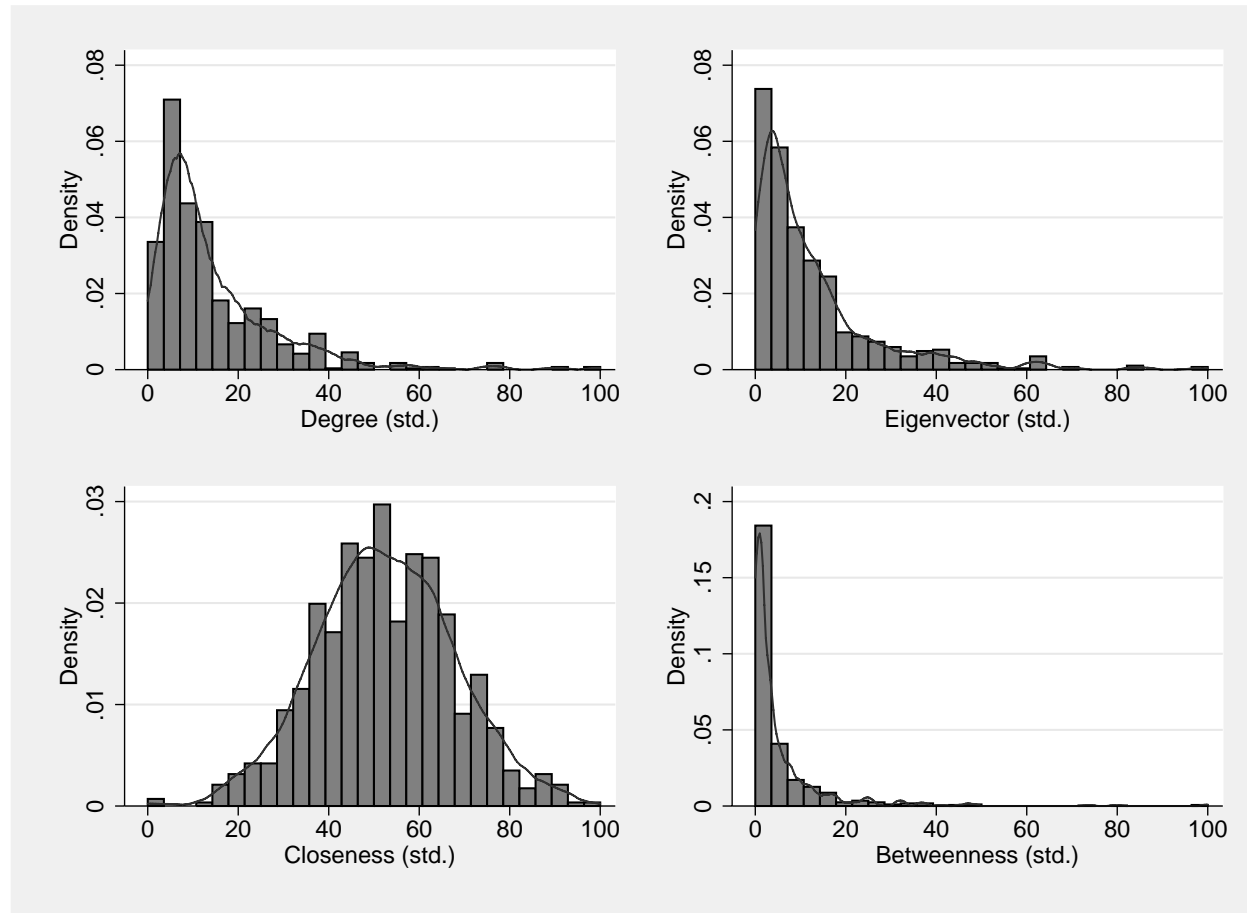


Figure 5: Densities of Centrality Indices

Figure 6 shows that all centrality measures grow steadily with age until age 60 and later start decreasing. Like typical earning profiles there is an inverse U-shaped relationship between centrality and age, but while for earnings profiles the peak is typically around 45 or 50, in the Mafia it is at least 10 years later. This finding confirms Ianni and Reuss-Ianni (1972, pg.130)'s anthropologic results about the importance of age in determining the leadership positions. The difference between the minimum age (23) and age (60) doubles the degree, and the eigenvector index, and triplicates *betweenness*. Closeness, instead, shows a very steep increase up to age 30 and then flattens out, but the overall increase is more modest.

Figure 7 presents how the interaction index, the 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.

The evidence collected so far is thus in line with some conjectured organizational rules of the Mafia, it shows that network centrality indicators adequately capture the importance of mafia leader (with different nuances) and, most importantly, it clearly documents that links are not random. As pointed out by Jackson and Rogers (2007), the distribution of *degree* in a socially generated network is more unequal than in random networks. In random networks the log-frequency is linear in log-*degree*. Figure 8 reveals that, in fact, this is not true for the Mafia network. 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 table shows positive and significant correlation, called

²⁰The individual centrality measures have been normalized to be between 1 and 100

Mastrobuoni and Patacchini: Organized Crime Networks

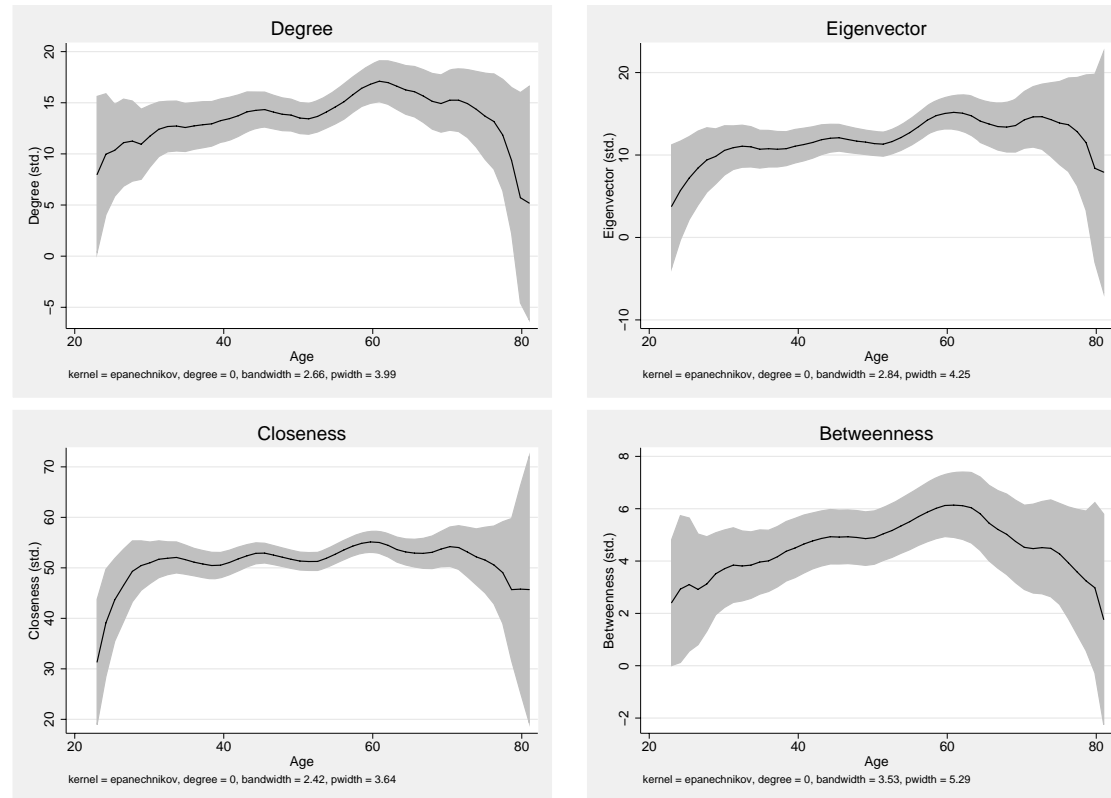


Figure 6: Centrality and Age

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

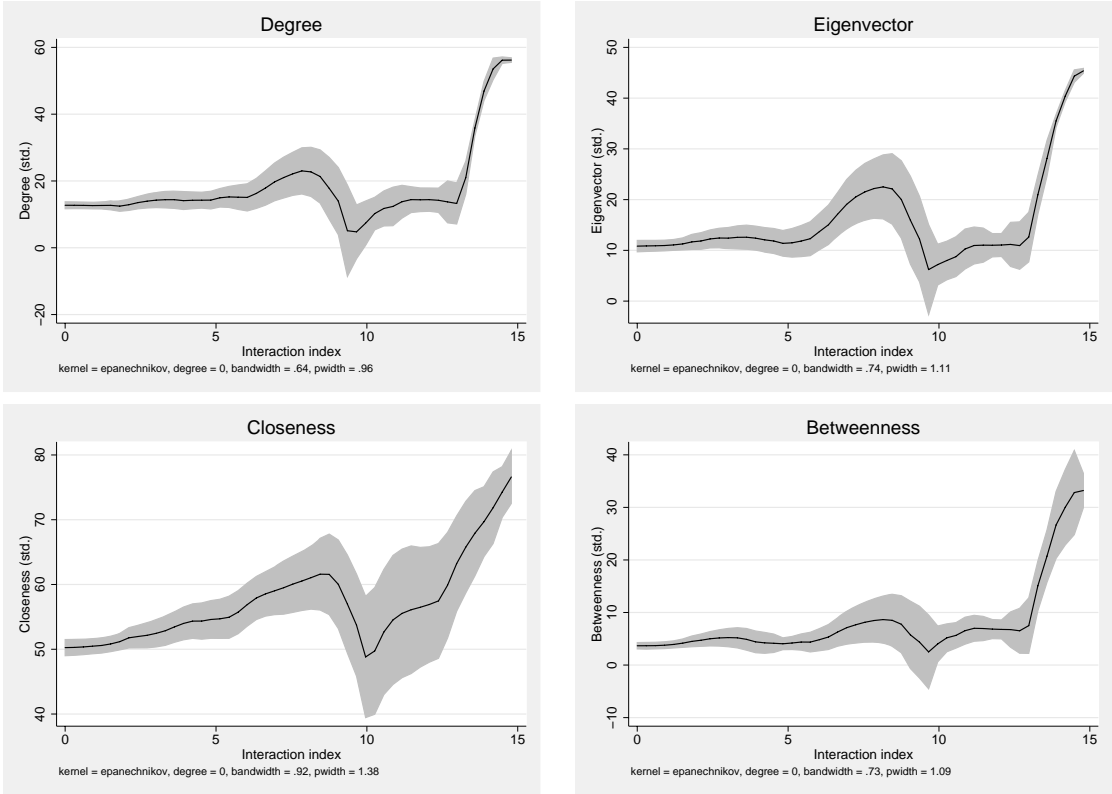


Figure 7: Centrality and “Tradition”

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

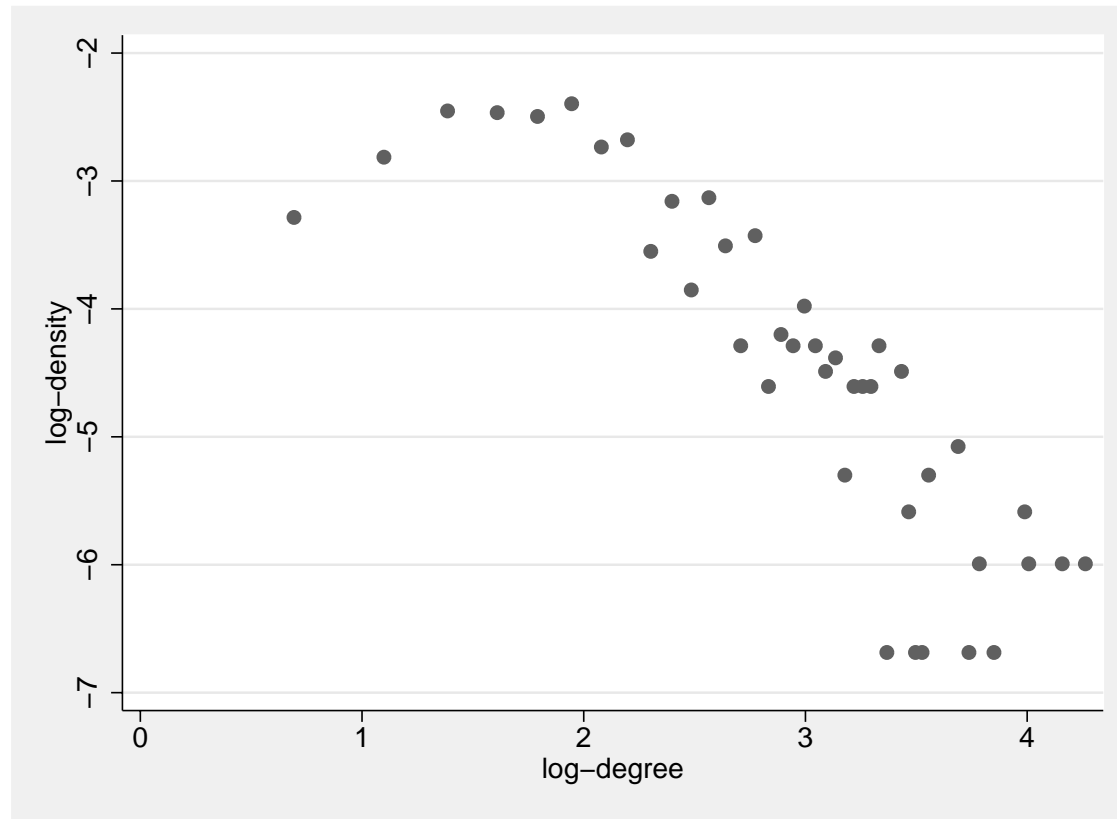


Figure 8: Log Density on log *Degree*

assortativity, for almost all crimes and legal businesses detailed in our data.

Table 7: Positive Assortativity Among Crime and Business Types

	Crimes		Businesses	
	$\hat{\beta}$	SD($\hat{\beta}$)	$\hat{\beta}$	SD($\hat{\beta}$)
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. I restrict the data to businesses held and crimes perpetrated by at least five percent of the sample.

The next section is going to describe mobsters depending on their centrality.

5 Who is a mafia leader?

We investigate the characteristics that correlate with our indicators of leadership using a simple linear relationship between the k -th index of importance c_i^k of individual i and his observable characteristics X_i : $c_i^k = \beta' X_i + e_i$, $k = 1, \dots, 4$, where e_i denotes the random error term.

Each model controls for the number of mobsters that in the 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 the connections are based on surnames, we add this variable to control for the mechanical effect that an increase in the number of mobsters might have on the centrality measures.

The Eigenvector Index

We start our simple regression analysis using the eigenvector index as dependent variable, as it depends both on the number and the quality of each mobster's connections (i.e. it takes into account both direct and indirect connections).

The first column of Table 8 contains the estimation results when using a parsimonious set of explanatory variables. Age enters quadratically. Peak centrality is reached at age 59 and both terms are significant. The finding of a peak late in life is consistent with Ianni and Reuss-Ianni (1972)'s account of hierarchies based on generations (though it might also reflect that more able mobsters are more likely to be alive). Another variable that has a significant positive influence on centrality is being born in Sicily (a 20 percent increase).

This means that not only is the American Mafia an Italian enterprise, it values 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. This is why we use the interaction index, which depends on the geographical distribution of last names, to proxy for community ties and exposure to the "Tradition." The index has a positive and significant effect on network centrality. Members from Sicily and those with stronger ties to the Mafia culture are likely to be trusted more, as they are more likely to adhere to the *omertà*. Body weight of the mobster does not influence his status.

In column 2 we control for additional factors, like the nuclear family structure of the mobster, that might in part be endogenous. The purpose is thus merely descriptive, and useful for predictive purposes. Each additional child increases the eigenvector index by 0.8. Children thus seem to be more valuable for higher ranked mobsters. 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 repre-

Table 8: Correlates of the Eigenvector Index

	(1)	(2)
	Eigenvector index	
Extended family members	4.672*** (0.905)	5.041*** (0.764)
Born in Italy (except Sicily)	-1.042 (1.128)	-0.599 (1.056)
Born in Sicily	2.578* (1.466)	3.387** (1.381)
Age	0.899*** (0.304)	0.909*** (0.315)
Age squared/100	-0.765** (0.300)	-0.768** (0.311)
Interaction index	2.512* (1.407)	2.388* (1.316)
Height in feet	2.049 (2.580)	2.534 (2.423)
Weight in pounds	0.001 (0.020)	-0.004 (0.018)
Married		-0.737 (1.057)
Divorced		-0.507 (2.369)
Connected wife		4.505*** (1.546)
Number of children		0.808** (0.347)
Fraction of daughters		1.778 (1.903)
Siblings		-0.115 (0.218)
Resides in NY		5.297*** (1.123)
Resides in NJ		2.734* (1.650)
Violent crimes		1.337 (0.931)
Age at first arrest		-0.057* (0.034)
Never arrested		-2.176* (1.199)
Types of crime committed		0.773** (0.299)
Types of businesses		1.915*** (0.527)
Observations	801	801
R-squared	0.195	0.287
Mean value:	12.57	12.57

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.

sents potential “workforce.” The 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 adopted a more liberal view towards divorce. Being married to a wife who is connected, instead, is associated to an almost 50 percent increase in the index. These findings, again, seem to suggest that having more trusted links is associated with an higher number of connections. However, 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 findings. The leading mobsters are active in New York and New Jersey. 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 combined regressors explain almost 30 percent of the variability of the eigenvector index. Criminals who were known to have committed violent crimes have a larger index, but the effect is not significant. An arrest increases the importance, but reverse causality might imply that higher ranked individuals are 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), but this effect is also not significant once we control for all the other factors.

Other Network Centrality Indices

As mentioned before, 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

Table 9: Correlates of Other Centrality Measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Degree	Eigenvector		Closeness		Betweenness	
		Total	Residual	Total	Residual	Total	Residual
Extended family members	5.631*** (0.873)	5.041*** (0.764)	-0.152 (0.494)	6.524*** (0.751)	2.844*** (0.585)	1.616*** (0.590)	-0.009 (0.037)
Born in Italy (except Sicily)	-0.363 (0.993)	-0.599 (1.056)	-0.265 (0.621)	-2.216* (1.259)	-0.407 (0.615)	-1.900** (0.956)	0.039 (0.093)
Born in Sicily	2.592** (1.318)	3.387** (1.381)	0.996 (0.682)	4.099*** (1.346)	1.642* (0.887)	1.840* (0.961)	0.153 (0.099)
Age	1.123*** (0.312)	0.909*** (0.315)	-0.126 (0.190)	1.386*** (0.368)	0.804*** (0.226)	0.407 (0.305)	0.015 (0.027)
Age squared/100	-1.010*** (0.307)	-0.768** (0.311)	0.163 (0.185)	-1.244*** (0.362)	-0.763*** (0.224)	-0.364 (0.299)	-0.002 (0.026)
Interaction index	2.783** (1.278)	2.388* (1.316)	-0.179 (0.797)	2.428*** (0.907)	3.582*** (0.841)	0.002 (0.820)	-0.056 (0.058)
Height in feet	3.180 (2.359)	2.534 (2.423)	-0.399 (1.230)	5.166* (2.677)	1.258 (1.615)	2.395 (1.818)	-0.104 (0.206)
Weight in pounds	0.003 (0.018)	-0.004 (0.018)	-0.007 (0.008)	-0.007 (0.021)	0.001 (0.013)	-0.009 (0.016)	0.001 (0.001)
Married	0.028 (1.006)	-0.737 (1.057)	-0.763 (0.538)	0.423 (1.126)	-0.535 (0.728)	0.399 (0.823)	-0.051 (0.087)
Divorced	-1.423 (1.973)	-0.507 (2.369)	0.805 (1.135)	-1.264 (2.693)	-0.511 (1.182)	-0.024 (1.984)	0.083 (0.163)
Connected wife	4.388*** (1.418)	4.505*** (1.546)	0.457 (0.734)	3.236** (1.315)	1.916** (0.910)	-0.589 (0.937)	0.107 (0.104)
Number of children	0.826** (0.336)	0.808** (0.347)	0.046 (0.178)	0.810** (0.326)	0.531** (0.255)	0.090 (0.239)	0.012 (0.029)
Fraction of daughters	0.552 (1.604)	1.778 (1.903)	1.269 (0.959)	0.944 (1.745)	-0.366 (1.091)	0.463 (1.265)	0.167 (0.142)
Siblings	0.113 (0.196)	-0.115 (0.218)	-0.219* (0.113)	0.224 (0.219)	0.184 (0.136)	0.126 (0.157)	0.037* (0.019)
Resides in NY	2.186** (0.963)	5.297*** (1.123)	3.281*** (0.612)	7.412*** (1.064)	0.379 (0.645)	5.507*** (0.761)	0.166** (0.084)
Resides in NJ	0.554 (1.916)	2.734* (1.650)	2.223** (0.914)	8.119*** (2.016)	1.839 (1.836)	7.637*** (1.633)	0.323* (0.189)
Violent crimes	0.733 (0.874)	1.337 (0.931)	0.662 (0.526)	0.602 (0.929)	0.561 (0.629)	-0.036 (0.708)	0.084 (0.082)
Age at first arrest	-0.014 (0.032)	-0.057* (0.034)	-0.043* (0.022)	-0.016 (0.041)	-0.007 (0.021)	-0.003 (0.033)	-0.002 (0.003)
Never arrested	-1.704 (1.141)	-2.176* (1.199)	-0.604 (0.692)	-1.312 (1.470)	-0.868 (0.763)	0.173 (1.098)	-0.053 (0.112)
Types of crime committed	0.668** (0.270)	0.773** (0.299)	0.157 (0.145)	0.568 (0.350)	0.259 (0.164)	-0.014 (0.230)	0.014 (0.024)
Types of businesses	1.406*** (0.475)	1.915*** (0.527)	0.619*** (0.236)	1.849*** (0.511)	0.796** (0.323)	0.623* (0.348)	0.087** (0.040)
Observations	801	801	801	801	801	801	801
R-squared	0.323	0.287	0.091	0.367	0.266	0.136	0.071

Notes: The regressions include also missing dummies 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

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 use a simple statistical way to accomplish the same goal: we simply consider 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 ease 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. A 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 links, while for all the other variables that are significant in column 2 but not in column 3 only direct links matter (connected wife, age, and the interaction index).

Closeness, an inverse measure of the average distance from the other members, and *Betweenness*, the ability to build bridges, capture instead different aspects in the definition of individual importance or power within a network. Column 4 shows that *Closeness* is 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 is driven by the indirect links (cannot be explained by *degree* alone).

Results that use *betweenness* as a dependent variable shown in column 6 present a 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 also have 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.

Qualitative Measures of Importance

In column 2 of Table 10 we exploit the qualitative information on the mobsters contained in the records. In particular, we use the same specification as before but using *Top Ranked Citations* defined in Section 4 as the dependent variable. 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 children, having strong community ties (interactions), and the number of crimes committed are no longer significantly related to leadership. This lack of significance leads to a lower R-squared (8 versus 28 percent). While there might be superior ways to extract the information known by the FBN (we did try several qualitative measures), network based measures seem to capture additional attributes of leadership.

6 Robustness checks

Table 11 shows three types of robustness checks. In the first (columns 1 and 2), instead of using undirected links we exploit the directed nature of the nominations, meaning that we define mobster i to be connected to mobster j if i appears as an associate in the criminal profiles of j . We thus use here only the *indegree*, i.e. the number of times someone appears as an associate in criminal profiles of other mobsters (standardized to lie between 0 and 100). Given that *outdegree* is somehow bounded, it is not surprising that the coefficients are driven by *indegree*, and thus are very close to the ones based on undirected connections shown in the first column of Table 9. The second robustness check relates to the truncation of the sample. Since we have information 800 out of an estimated 5,000 “made” men across the U.S., our sample is truncated. We attempt to simulate the possible bias of our results that a truncation might generate (see Borgatti et al., 2006, for a similar simulation). 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. Truncation tends to reduce the size of the coefficient on the dummy “connected wife,” which is reasonable given that smaller samples lead

Table 10: Correlates of the Qualitative Importance Indicators

	(1) Eigenvector	(2) Top Ranked
Extended family members	5.041***	-0.009
Born in Italy (except Sicily)	-0.599	0.039
Born in Sicily	3.387**	0.153
Age	0.909***	0.015
Age squared/100	-0.768**	-0.002
Interaction index	2.388*	-0.056
Height in feet	2.534	-0.104
Weight in pounds	-0.004	0.001
Married	-0.737	-0.051
Divorced	-0.507	0.083
Connected wife	4.505***	0.107
Number of children	0.808**	0.012
Fraction of daughters	1.778	0.167
Siblings	-0.115	0.037*
Resides in NY	5.297***	0.166**
Resides in NJ	2.734*	0.323*
Violent crimes	1.337	0.084
Age at first arrest	-0.057*	-0.002
Never arrested	-2.176*	-0.053
Types of crime committed	0.773**	0.014
Types of businesses	1.915***	0.087**
Observations	801	801
R-squared	0.287	0.071

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 height and weight. Clustered (by family) standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

to an increase in misclassification error of that particular variable. The other coefficients tend to be quite stable.

Table 11: Robustness Checks

	β	$se(beta)$	simulated		β	$se(beta)$
			average	median	MC	weighted
Extended family members	5.541***	0.861	6.031	5.916	3.991***	0.737
Born in Italy (except Sicily)	-0.437	0.975	-0.294	-0.262	-0.264	0.663
Born in Sicily	2.577**	1.293	2.968	3.082	2.183***	0.814
Age	1.111***	0.308	1.148	1.139	0.315*	0.178
Age squared/100	-0.999***	0.302	-1.03	-1.023	-0.243	0.177
Interaction index	2.766**	1.262	3.023	2.764	2.363**	1.099
Height in feet	3.094	2.326	3.066	3.24	2.137	1.483
Weight in pounds	0.003	0.018	0.005	0.005	0.004	0.009
Married	0.041	0.988	0.628	0.539	0.288	0.596
Divorced	-1.54	1.955	-0.578	-0.602	-1.137	1.4
Connected wife	4.777***	1.5	4.277	4.018	2.925***	0.861
Number of children	0.810**	0.332	0.821	0.802	0.15	0.212
Fraction of daughters	0.531	1.576	0.408	0.453	1.022	1.087
Siblings	0.124	0.194	0.088	0.095	-0.093	0.132
Resides in NY	2.114**	0.947	2.532	2.563	3.485***	0.626
Resides in NJ	0.69	1.893	0.71	0.651	2.376*	1.346
Violent crimes	0.692	0.861	0.631	0.558	-0.036	0.492
Age at first arrest	-0.018	0.031	-0.023	-0.023	-0.032	0.023
Never arrested	-1.703	1.123	-1.779	-1.696	-0.741	0.794
Types of crime committed	0.645**	0.267	0.673	0.671	0.219	0.203
Types of businesses	1.344***	0.466	1.481	1.518	0.666**	0.283

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.

In the third robustness check we address more in detail the non-random nature of our data by assigning to each observation a weight that is proportional to the probability of being included in the sample. The more a mobster is connected, the more he becomes exposed, and likely to be monitored and, thus recorded in our data files. The adjacency matrix of a network \mathbf{G} keeps track of the direct connections, the k th power $\mathbf{G}^k = \mathbf{G}^{(k \text{ times})} \mathbf{G}$ of this ad-

jacency matrix then keeps track of indirect connections. Such matrices can be used to detect the probability that, given an initial sample of network members, each other member is going to be detected. This approach resembles a sampling procedure that has been called Respondent-Driven-Sampling (RDS) (Heckathorn, 1997). Indeed, there are no exact records about how the FBN followed mobsters and constructed the network, though with the use of surveillance posts, undercover agents, etc. agents were probably discovering previously unknown mobsters following known ones.

If one starts with an initial fraction of P mobsters that are observed with probabilities p_0 , that is a $1 \times N$ vector of zeros and ones (if observed) row normalized, called the seed, after k steps the likelihood of observing a mobsters is

$$p_0 G^k = p_k$$

where the row-normalized G in this context indicates the transition matrix. The solution of such a system,²¹ i.e. the stationary distribution p , does not depend on the seed and represents the the likelihood that a mobsters has been observed after several steps.

The corresponding resampling weights are going to be simply $w_i^0 = \frac{1}{p_i}$, with $0 < p_i < 1$, where p_i is the generic element of the eigenvector associated to the largest eigenvalue of G which is equal to 1 if the matrix is row standardized.

The last two columns of Table 11 contain our regression results (9) with weighted observations. The weighting scheme lowers the average of the eigenvector centrality index from 12.57 to 6.93. With this qualification in mind, it appears that the results remain basically unchanged, i.e. the relative variation with respect to the sample means is unaffected by the weighting scheme.

7 Concluding remarks

This paper presents the first through micro-level analysis of the US Mafia network. A network-based analysis of this structure allows us to find empirical support of sociological and historical views about the functioning of these criminal networks, with notable qualifications. Unlike economic organizations hierarchies depend crucially on kinship. We document that variables that

²¹The Perron-Frobenius theorem ensures that such a vector exists and is unique.

lower the risk of defection – family ties, violence, and mafia *culture* among others – increase the number of connections. In particular, we find signs of strategic endogamy. Our results are consistent with the view that women are used to foster the Family’s network centrality. However, this is only true within trusted Families. Indeed, our results reveal that women are not used to bridge Families that are not otherwise closely connected. We also find that 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 have more businesses, legal and illegal ones.

Mathematical modeling that ties social network structure with 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 our 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 reveals some relevant features of social structure that can guide crime prevention policies. A targeted policy identifying “key players” in a given area may be an effective way to reduce crime (Ballester et al., 2006, Réka et al., 2000). A necessary condition for designing such policies is the ability to map and identify a social network structure. Detailed information about the hierarchy might even allow enforcement agents to break the chain of command by arresting sets of mobsters (Farley, 2003). Only detailed empirical studies on real-world social networks of criminals can provide guidance in this direction.

A Centrality Measures

Let $N = \{1, \dots, n\}$ be a finite set of agents in network \mathbf{g} . Let us define \mathbf{G} as the n -square adjacency matrix of the 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.

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

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

This definition of centrality is thus based on the number of direct links only. A variant of 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 the 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 .²² 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*

²²The length of a shortest path is the smallest k such that there is at least one path of length k from i to j . I can identify such a length by computing \mathbf{G} , \mathbf{G}^2 , \mathbf{G}^3 , ..., until I find the first k such that the (i, j) th entry of \mathbf{G}^k is not zero.

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

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

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 .

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.
- Wayne E. Baker and Robert R. Faulkner. The Social Organization of Conspiracy: Illegal Networks in the Heavy Electrical Equipment Industry. *American Sociological Review*, 58(6):pp. 837–860, 1993.
- C. Ballester, A. Calvó-Armengol, and Y. 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.
- Stephen P. Borgatti, Kathleen M. Carley, and David Krackhardt. On the Robustness of Centrality Measures Under Conditions of Imperfect Data. *Social Networks*, 28(2):124–136, 2006.

- 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.
- Nigel Coles. It's Not What You Know-It's Who You Know That Counts. Analysing Serious Crime Groups as Social Networks. *British Journal of Criminology*, 41(4):580, 2001.
- David Critchley. *The Origin of Organized Crime in America: The New York City Mafia, 1891–1931*. Routledge, 2009.
- Scott H. Decker, Tim Bynum, and Deborah Weisel. A Tale of two Cities: Gangs as Organized Crime Groups. *Justice Quarterly*, 15(3):395–425, 1998.
- Giovanni Falcone and Marcelle Padovani. *Cose di Cosa Nostra*. Rizzoli, 1991.
- Jonathan D. Farley. Breaking Al Qaeda cells: A mathematical analysis of counterterrorism operations (A guide for risk assessment and decision making). *Studies in Conflict & Terrorism*, 26(6):399–411, 2003.
- 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.
- 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.

- Douglas D. Heckathorn. Respondent-driven sampling: a new approach to the study of hidden populations. *Social problems*, pages 174–199, 1997.
- 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.
- Peter Klerks. The Network Paradigm Applied to Criminal Organisations. *Transnational Organised Crime: Perspectives on Global Security*, page 97, 2003.
- Valdis E. Krebs. Mapping Networks of Terrorist Cells. *Connections*, 24(3): 43–52, 2002.
- 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.
- Peter Maas. *The Valachi Papers*. Putnam, New York, 1968.
- Jean Marie McGloin. Policy And Intervention Considerations of a Network Analysis of Street Gangs. *Criminology & Public Policy*, 4(3):607–635, 2005.
- 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.
- Mangai Natarajan. Understanding the Structure of a Drug Trafficking Organization: A Conversational Analysis. *Crime Prevention Studies*, 11:273–298, 2000.
- Mangai Natarajan. Understanding the Structure of a Large Heroin Distribution Network: A Quantitative Analysis of Qualitative Data. *Journal of Quantitative Criminology*, 22(2):171–192, 2006.
- 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. *Annual Review of Sociology*, 24:1–24, 1998.
- Albert Réka, Hawoong Jeong, and Albert-László Barabási. Error and Attack Tolerance of Complex Networks. *Nature*, 406:378–382, 2000.
- Peter Reuter. *Disorganized Crime: the Economics of the Visible Hand*. MIT press, 1983.
- Peter Reuter. *Handbook of organized crime in the United States*, chapter Research on American Organized Crime. Greenwood Press, Westport, CT, 1994.
- Jerzy Sarnecki. Delinquent Networks in Sweden. *Journal of Quantitative Criminology*, 6(1):31–50, 1990. ISSN 0748-4518.
- Jerzy Sarnecki. *Delinquent Networks: Youth Co-offending in Stockholm*. Cambridge Univ Pr, 2001. ISBN 0521802393.
- 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.
- 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.
- Phil Williams. Transnational Criminal Networks. *Networks and netwars: the future of terror, crime, and militancy*, page 61, 2001.
- George Wolf and Joseph DiMona. *Frank Costello: Prime Minister of the Underworld*. Morrow, 1974.