Returns to Education in Criminal Organizations: Did Going to College Help Michael Corleone?*

Nadia Campaniello†, Rowena Gray‡, Giovanni Mastrobuoni§

February 2016¶

Abstract

Is there any return to education in criminal activities? This paper is one of the first to investigate whether education has not only a positive impact on legitimate, but also on illegitimate activities. We use as a case study one of the longest running criminal corporations in history: the Italian-American mafia. Its most successful members were capable businessmen, orchestrating crimes that required abilities that might be learned at school: extracting the optimal rent when setting up a racket, weighting interests against default risk when starting a loan sharking business or organizing supply chains, logistics and distribution when setting up a drug dealing system. We address this question by comparing mobsters to a variety of samples drawn from the United States 1940 Population Census, including a sample of their closest (non-mobster) neighbors. We document that mobsters have one year less education than their neighbors on average. We find that mobsters have significant returns to education of 7.5-8.5 percent, which is only slightly smaller than their neighbors and 2-5 percentage points smaller than for U.S.-born men or male citizens. Mobster returns were consistently about twice as large as a sample of Italian immigrants or immigrants from all origin countries. Within that, those charged with complex crimes including embezzlement and bookmaking have the highest returns. We conclude that private returns to education exist even in the illegal activities characterized by a certain degree of complexity as in the case of organized crime in mid-twentieth century United States.

*We acknowledge the excellent research assistance of Patryk Bronka, Leo Faiola, and Gautham Udupa. We would also like to thank the editor Lance Lochner for all his comments and Dan Hamermesh for suggesting a more catchy title.

†Department of Economics, University of Essex and IZA Email: ncampa@essex.ac.uk.

‡Department of Economics, University of California at Merced, Email: rgray6@ucmerced.edu.

§Department of Economics, University of Essex, Collegio Carlo Alberto, and IZA, Email: gmastrob@essex.ac.uk.

Keywords: Returns to education; organized crime; mafia; Italian-American Immigration; Federal Bureau of Narcotics; 1940 Census
JEL classification codes: I26; N32; K42; Z13

1 Introduction

Additional years of education are known to increase earnings in legitimate labor activities. But, what about illegal ones? In this study we will not discuss the activities of common criminals. Our focus is professional criminals who belonged to one of the most successful and long-lasting criminal organizations: the Italian-American mafia between the 1930s and the 1960s. We match a list set up by the Federal Bureau of Narcotics (FBN) of 712 mobsters belonging to the Italian-American mafia with the 1940 United States (U.S.) Census of Population. This gives us information about income, housing values, education, job characteristics, as well as the precise address of residence for each individual. We create a sample of white, male, similarly aged, neighbors of these mobsters that serves as the closest comparison group and we also present estimated returns to education for other samples drawn from the 1940 Census, including all working-age white men; whites born in the U.S.; all U.S. citizens; all immigrants; all Italian immigrants; and second-generation Italians (who are born in the U.S. but have at least one parent born in Italy).

Economists have shown that increased levels of education reduce criminal participation. This implies that education is valued more by legitimate firms than by illegitimate ones. This is consistent with our first finding: mafia mobsters have on average one less year of education when compared to the sample of neighbors.

But, this finding does not imply that annualized returns to education are smaller for organized crime members than for ordinary workers. Criminal careers are known to start very early and are likely to be interwoven with schooling choices. Individuals who choose to be part of the mafia are likely to trade off income and power for risk of injury, prison, and death.

This alone, without the need of lower returns to education, would predict a lower investment in education as there would be fewer years of working life in which to recoup foregone wages (Mincer, 1974). Indeed, economic theory predicts that individuals with *lower* (working) life expectancy should have *larger* annualized returns to education.

This is true unless the extra schooling is not marketable. So, is schooling marketable in the mafia? This question really involves the mafia's complex business model and the link between human capital and schooling. Let us start with the latter. If one takes Bowles and Gintis (2002)'s view that schools "prepare people for adult work rules, by socializing people to function well (and without complaint) in the hierarchical structure of the modern corporation" it would seem that schools are an ideal training environment for aspiring mobsters.

While we do not fully embrace this view of schooling, many of the skills students acquire at school are likely to be useful when setting up a racket (i.e. extracting the optimal rent), a loan sharking business (i.e. weighting interests against default risks), a drug dealing system (i.e. setting up supply chains), etc. It is ultimately an empirical question as to whether the returns to education in the mafia are similar to the ones ordinary workers enjoy. This comparison, we believe, is also informative about the workings of the mafia. The results presented below hold, we argue, for criminals engaged in complex criminal activities, but may not be more generally true of petty criminals or criminals operating at the lowest levels of criminal organizations, whose everyday tasks are much simpler and do not involve the planning, risk evaluation, and communication skills needed of those higher in the network. We are therefore providing a counterpoint to Carvalho and Soares (forthcoming), and Levitt and Venkatesh (2000), who study the characteristics of regular gang members.

We estimate Mincer-type regressions using log income and log housing value as the main outcomes. The main independent variable of interest is years of education. We present results for the mobster sample and compare to other reasonable comparison groups: all men; U.S. citizens; immigrants; Italian immigrants and second-generation Italian men; and a sample of mobster neighbors, who lived on the same block (and usually the same exact street) in 1940.

We find large returns to education within the mafia, no matter the model, or the outcome variable, that we use. This shows that private returns to education exist not only in legitimate but also in the illegitimate activities that imply a sufficient degree of complexity. Mobster returns (in terms of income) to a year of schooling are around 7.5-8.5 percent, compared to 9-10 percent for the neighbor sample and 10.5-13 percent for the U.S. born and U.S. citizen samples. Interestingly, mobster returns are substantially larger than we find for the immigrant and, especially, the Italian immigrant, samples, while they are only about one percentage point higher than we find for second-generation Italians. Moreover, for mobsters who, according to the FBN records, were involved in white-collar crimes or in crimes that require running an illegal business (i.e., racketeering, loan sharking, bootlegging, etc.) we find returns to education that are about three times as large as for those who are involved in violent crimes (i.e., robberies, murders, etc.).

To our knowledge, this is the first systematic attempt to estimate the returns to education in criminal activities and provides intuitive insights into the workings of complex criminal gangs such as the mafia and into the factors considered by those deciding to become criminals in the first place. Carvalho and Soares (forthcoming) provide some evidence on the returns to education for low level Rio de Janeiro gang members but it is not the main focus of their study.

The paper proceeds as follows. We first discuss the existing literature on education and crime, before providing a brief overview of the history of mafia organizations and members in the U.S. before 1960. We then present our novel dataset as well as our comparison samples drawn from the U.S. Census. We then discuss the empirical methodology before finally presenting our results, discussing mechanisms, and concluding.

2 Literature Review

This section discusses both the existing literature analyzing the impact of education on crime and the recent and historical literature measuring the returns to education, providing context to the analysis presented below.

Education has been relatively neglected by economists as a channel that might influence both criminal proceeds and the incentive to engage in crime. Most of the (quite recent) literature on this topic finds that education reduces crime through different potential channels: first of all, education increases the opportunity cost of the legitimate labor market and, if arrested, of the lost time spent in jail, thus making it more costly to engage in criminal activities ("deterrence" mechanism). Then, education might change preferences and behaviors, by increasing patience and risk aversion. Finally, there is a potential "school incapacitation" effect that might lead to a direct decrease in crime. Lochner and Moretti (2004), Machin et al. (2011) and Hjalmarsson et al. (2015) use a school reform as an exogenous source of variation for an extra compulsory year of schooling and find strong evidence that education significantly decreases the probability of being involved in criminal activities.

Meghir et al. (2012) find that the Swedish educational reform introduced in 1962 aimed at increasing the number of years of compulsory schooling had not only a direct effect in reducing crime for the men affected by the reform, but also for their children. Anderson (2014) uses county-level data for the U.S. to show that arrest rates for 16-18 year olds are significantly lowered by the implementation of minimum age of dropout laws. Luallen (2006) investigates the impact of school incapacitation on juvenile crime rates by using teacher strike days as a source of variation of student school attendance and he finds that schooling significantly decreases juvenile crime.[1].

And yet, Ehrlich (1975) suggests that the relation between education and crime may

[1]For a more comprehensive literature review of the effect of education on crime see Lochner (2011).

be more complex, since it depends on the way education affects the relative opportunities available to offenders in different illegitimate activities. In his view, education can be regarded as an instrument to improve efficiency in the production in both legitimate as well as illegitimate markets, and we should expect to find lower educated people committing petty crimes, and more educated ones committing more elaborate crimes (e.g. fraud, forgery, embezzlement, trade in illegal merchandise, and illegal commercial practices, etc). In addition, education may increase an offender's ability to avoid apprehension and punishment for their crimes. While Ehrlich (1975) examines data relying on an intuitive model, Lochner (2004) explicitly models the decision to invest in human capital, to work, and to commit crime. He adopts a human capital framework to explore the relationship between education and crime and his findings are in line with Ehrlich (1975)'s intuitions. The predictions of his model are generally supported by the data presented in the empirical part of his paper where he finds that education is associated with fewer property and violent crimes but with more white collar crimes (although not significantly).

Lack of individual data on criminal proceeds *and* education has prevented scholars from analyzing the effect of education and on the productivity of criminals.[2] The only exception we are aware of is Carvalho and Soares (forthcoming), a recent paper on 230 youngsters working for drug-selling gangs in 34 poor neighborhoods of Rio de Janeiro (so called "favelas"), Brazil. The authors have very detailed information on socioeconomic factors, like years of schooling, literacy, wages related to drug dealing, involvement in violence, etc. Their study is not focussed on estimating the returns to education but in their Mincer wage regressions the coefficient on years of education is not significantly different from zero. Instead, the coefficient on years of experience ranges between 5 and 10 percent.

[2]Moreover, the data typically used to study the relationship between criminal participation and education is based on prison records. Inmates might not be a representative sample of all criminals, but just a selection of the least able, thus underestimating the level of education and its return for common offenders.

But, drug selling in a Brazilian favela is likely to require a different set of skills compared to many of the legal and illegal businesses that were run by the mafia in New York and in other major U.S. cities historically. The involvement of victims in racketeering, extortions, and fraud adds an additional layer of complexity which is more common in white-collar crimes. Moreover, many of these businesses were often run together, again, adding complexity.

Levitt and Venkatesh (2000) investigate the characteristics of members of a gang located in an inner-city neighborhood in a large, industrial American city. They show that gangsters' average wages are only slightly higher than those earned in the legal sector, but that the distribution of wages is highly skewed and is characterized by enormous wage differentials between the gangsters at the bottom and those at the top of the criminal organization. They interpret the decision to join a gang as a tournament, where the winners will be highly compensated in terms of future wage. But, they have no data on the educational attainment of gangsters.

This paper also relates to the large literature estimating the private returns to education more generally. For several decades, economists have been running Mincer regressions similar in form to those we present and estimate below, variously using ordinary least squares (OLS), instrumental variables (IV) and control function techniques to address estimation issues including ability bias and measurement error. Recent investigations by Heckman et al. (2003) have found that the Mincer specification, which assumes a linear relationship between log earnings and years of education and a quadratic relationship between log earnings and experience, was most appropriate for the period 1940-1950. This is reassuring for the results presented in this paper and indicates that our estimates can reasonably be considered to represent the internal rate of return to education.

Ashenfelter et al. (1999) provides a meta-analysis of 27 modern studies estimating the returns to schooling, focusing mainly on twins and sibling studies where estimates are based on within-family variation, and on IV analyses. Returns based on OLS estimation of Mincer-type regressions tend to average 6-7%, while using IV or a twins sample yields estimates closer to 9% on average. Their method controls for reporting bias whereby studies finding insignificant results tend to be underreported, which may be a particular problem for IV and twin studies given their larger sampling errors. Once they employ this approach, they conclude that the estimated returns to schooling identified in the literature do not differ substantively due to differing estimation strategies. This conclusion is reassuring for us, given that we are limited due to the nature of our historical data in this study in terms of moving beyond OLS estimations.

Card (2001) surveys the current state of the literature, focusing on IV approaches. He points out that, even in studies using the most convincing instruments[3], the interpretation of the results must be as the average effect of education on earnings across individuals with potentially heterogeneous returns to and costs of obtaining education and it also reflects who was most affected by the instrument, which may not always be representative of the returns to education of the average person in the population. Given that the returns may be higher for those at lower levels of education, and that most IV strategies tend to exploit this margin of exogeneity in attainment (the compulsory schooling and distance to educational institutions studies for example), it is not so surprising that IV estimates tend to be larger. This also suggests that producing OLS estimates is still a useful exercise.

Finally, we discuss the smaller literature on education and estimates of its return in pre-World War II U.S. The historical literature on education has traditionally focused on plotting the general trend of the rise of educational attainment and public education.[4] The general trend during the early twentieth century was a steep upward trajectory in educational attainment associated with the "High School Revolution", with some states

[3]And there is evidence suggesting that some of those studies have used weak instruments, including quarter of birth, which would bias the estimated coefficients towards OLS, Card (1999), p. 1837.

[4]See, for example, the large body of work by C. Goldin and L. Katz, including Goldin and Katz (2008a) and Goldin and Katz (2008b).

in New England and the Midwest increasing attainment faster than others. By 1940, half of U.S. youths had attained a high school diploma (Goldin and Katz (2000), p. 786).

Lack of data on wages or income and on educational attainment before 1940 has held back estimation of the returns to education for earlier dates. Clay et al. (2012) comprehensively examine the returns to education over the long run during the period when compulsory schooling laws were first introduced on a state by state basis and therefore when they were most relevant. They look at men reporting positive wage income in the 1940 Census and who were impacted by the laws of 1898-1927 and estimate returns to an additional year of education of 8-9%, using OLS, and 11-14% using IV methods where the compulsory schooling laws provide a plausible instrument. Heckman et al. (2003) found that the returns to education in a sample of white men aged 16-64 drawn from the 1940 Census were about 12.5%. Our estimated returns for the comparison groups drawn from the U.S. Census, including the sample of neighbors, reported below, are very much in line with existing OLS estimates from the historical literature.

3 The Italian-American Mafia

This section provides some context regarding the Italian-American population and mafia from the turn of the twentieth century onwards, which will inform our analyses of rates of educational attainment and measured returns to schooling for these groups.

From 1880 to 1900, 959,000 Italians entered the U.S., and the following two decades saw a further 3,200,000 Italians make the journey.[5] This massive wave of migration to the U.S. stopped with World War I and the introduction of immigration restriction in the 1920s and, particularly, the Immigration Act of 1924, after which the annual visa quota for Italians was reduced to 4000. While initially the source locations were found in Northern Italy, over time Sicily and the South provided a larger proportion of new arrivals, due

[5] Pretelli (2013), p. 437.

to labor unrest, population excess and, most of all, agricultural crises and commodity price shocks.[6] Buonanno et al. (2015) have shown how these agricultural problems also contributed to the rise of the mafia in many parts of Sicily and Southern Italy.

The majority of these immigrants were agricultural workers, with low levels of literacy. In 1901, about 80 percent of the Sicilian population was illiterate (ISTAT, 2014), and such rates were likely similar among the negatively selected group of early immigrants (immigrants tend to be younger and thus more literate but also poorer and thus less literate). The early immigrants tended to be geographically clustered, with large numbers living in little "Italies" and, as a group, they maintained their hostility to schooling. The "Americanization" that might occur in American public schools was perceived as a threat to their values. For example, Anthony Accardo, Chicago's boss-of-bosses for almost a half-century (who has a record in the FBN data), was born in Chicago to Sicilian immigrants. Both settled in the U.S. in 1905. When Anthony was 14 his parents filed paperwork with the authorities claiming that he was two years older than he actually was so that he could leave school and go to work, a common practice (Roemer, 1995). Later, we show that it is precisely after eight years of schooling, when children are about 14, that the educational gap between mafia members and their neighbors emerges.

Possibly also because of the educational gap, children of these early immigrants had a higher tendency to become street gang members in the slums, spoke little Italian, and worked side by side with criminals from other ethnicities, mainly Jewish and Irish (Lupo, 2009). Several mafia bosses, including Lucky Luciano, Tommaso Lucchese, Vito Genovese, Frank Costello, etc., were children of these early immigrants. Criminal careers started quite early– FBN records show that in fifty percent of cases the very first recorded arrest occurred before the age of 20.

Lured by the criminal successes of the first wave of immigrants, and by Prohibition, the

[6] Between 1901 and 1913, almost a quarter of Sicily's population departed for the United States, Critchley (2009).

second wave of immigrants that went on to become mafia bosses were already criminals by the time they entered the U.S. Charles Gambino, Joe Profaci and others were in their 20s and 30s when they first entered the U.S., and most came from Sicily. Another reason for this selection of immigrants was the 1920s fascist crack-down of the mafia, which forced some of these criminals to leave Sicily. After the second wave of immigration, the Italian-American mafia became more closely linked to the Sicilian mafia.[7] Lucky Luciano became the head of the entire organization in 1931 and there would be no more mafia wars during his reign which lasted until 1962– this decreases the problem of sample selection due to death in our mobster sample.

By 1940, the mafia had a well established government, called "commissione." The mafia, the *Cosa Nostra* ("our thing"), was composed of approximately 25 Families and was governed by a *Commissione* of 7-12 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. Each Family was structured in hierarchies with a boss (*Capo Famiglia*) at the top (Maas, 1968). These hierarchies allowed the mafia to successfully expand into a series of legal and illegal activities. Mobsters were involved in racketeering,[8] drug trafficking, gambling and bootlegging, but also owned restaurants, drugstores or were otherwise involved in the food and beverage sector. Real estate, casinos, car dealerships, loan-sharking and import-export were also common businesses. According to the FBN files, by 1960 only 32 percent of gangsters had no businesses, while 43 percent had one, 19 percent had two, and the remaining 5 percent had 3, 4, or 5 different businesses.

So, it seems clear that our sample mobsters represent career criminals engaged in elaborate crimes requiring a complex hierarchy of individuals. Ferrante (2011), a former member of the mafia associated with the Gambino family, describes these types as follows:

[7] See Gosch and Hammer (1975).

[8] Gambetta (1996) views the mafia as a protection agency that in exchange of a fee allows firms to collude. Alexander (1997) shows evidence of such collusion practices in the 1930 Chicago Pasta market.

> "If we shed our prejudices, we'll find that accomplished mobsters are just like top business leaders. The mafia shares the same power structure as any corporation. A Don is exactly like a CEO, steering the business (or family) into the future. His capos are middle-managers or department heads, and his soldiers are employees. Whether corporate or mafia, people who acquire diplomatic skills, leadership qualities, and the enthusiasm to motivate will master their respective fields."

Later, we show that the returns to education for these "business" criminals are, indeed, quite large, and they drive our general finding of a healthy return to education for mobsters.

4 Data

4.1 Mobster and Neighbor Samples

In this section we explain how the dataset was constructed. We searched the manuscript records of the 1940 U.S. Census of Population for 712 members of the Italian-American mafia whose details were listed in the 1960 Federal Bureau of Narcotics (FBN) records.[9]

The records are an exact facsimile of the FBN secret files on American mafia members who were active in 1960 (FBN, 2007).[10] There are no written records about how the FBN followed mobsters and constructed the network. Through surveillance posts and undercover agents, the agency was likely discovering previously unknown mobsters by following known ones. Very active mobsters with many connections are thus likely to be

[9] In the 1930s, and up to 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). The New York Federal Bureau of Investigation had just a handful of agents assigned to the mafia, while in the same office more than 400 agents were fighting domestic communists (Maas, 1968).

[10] The distribution of the year of first arrest of mobsters has almost full support within the range 1908-1960, so one can infer that the data refer to what the authorities knew in 1960.

over-represented in our dataset. In some specifications, we use information on the network of connections to control for this sample selection, by reweighting the data in a way that resembles snowball sampling (see Mastrobuoni (2015)).

We then link these records based on a multitude of variables (first name, surname, names of family members, the residence address, the year of birth, etc.) *by hand* to the 1940 Decennial Census using the genealogical Web site ancestry.com. We face two selection problems. We can only match mobsters who survived up to 1960 and we can only gather information on incomes and/or housing outcomes when the mobsters are not in prison at the time of the Census. A smaller issue is the fact that some mobsters immigrated to the U.S. after 1940, but we find evidence that, for those we managed to find on "ancestry.com", this was only the case for 7 of the 712 potential mobsters, so we do not consider this to be a big problem.

Between 1940 and 1960, mafia Families were not at war with each other, so the second selection problem is likely to be more serious. Thirty-two mobsters out of the 414 (7.7 percent) that we matched to the 1940 Census were in prison. One would expect the more "executive" members of the mafia, the soldiers, to be more likely to face the risks of prison or death. We have information on education for those spending time in prison, and these mobsters do indeed have, on average, lower levels of education compared to the ones that are out of prison (6.8 versus 7.7 years).

Including the 32 inmates, by imputing their incomes or housing values, has little influence on the estimated returns to education, but we need to keep in mind that this robustness test cannot be performed for the members who died between 1940 and 1960.[11]

While we also cannot exclude that a few of these mobsters might not yet have been part of Cosa Nostra in 1940, this is unlikely to be serious problem. For each mobster the FBN record contains their criminal history, including the year of first arrest (only 15 percent of mobsters have no arrest record). For 85 percent of mobsters that we match

[11]The results are available upon request.

the first arrest happens in 1940 or before.[12]

Our search achieved a high match rate of almost 57%, matching 414 individuals to their Census record. This compares favorably with match rates from other studies searching for individuals in historical Censuses using ancestry.com– our relative success is likely attributed to the amount of information in the FBN records that we could use to match to the Census. Using broadly similar search criteria, Collins and Wanamaker (2013) obtain a 21% match rate when searching for African-American men between the 1910 and 1930 U.S. Censuses, while Abramitzky et al. (2014) report a 19% match rate when connecting more than one census between 1900 and 1920.

Table 1 compares the characteristics of our final sample (matched and selected based on age, etc.) with the unmatched or unselected mobsters. The purpose is to highlight any difference that might influence our results. From the FBN records, we compare the year of birth, the business criminals indicator, the number of words in the records that describe a mobster as a top ranked member of the mafia, marital status, the number of children, whether they resided in New Jersey or in New York and, finally, whether they had ever been arrested and, if that is the case, the year of their first arrest.

The number of words variable is constructed using more detail from the FBN files, which contain descriptions of each mobster and their activities. Our "Top ranked citations" variable sums up the number of mentions of the following words: boss, highest, most, head, top, high, influential, important, leader, leading, powerful, and representing.

We define "business" criminals as those who are charged with any of the following offenses: embezzlement, forgery, fraud, counterfeiting, gambling, prostitution, tax evasion, and bookmaking. "Non business" criminals are all the others. The most common "non business" crimes are: robbery, murder, weapons offenses, simple assault, larceny, burglary, drug and liquor offenses. Slightly more than one third of mobsters are business criminals.[13]

[12] As a robustness check we can restrict the analysis to these mobsters and the returns to education are almost identical for the two groups. Results are available upon request.

[13] Adding liquor and drug offenses to the definition of business criminals does not alter the results.

The averages tend to be close to each other. Unsurprisingly, given the 20-year differences between the two data sets, unmatched mobsters tend to be slightly younger (two years on average). Probably because of this difference they are also less likely to be married, and have on average fewer children. Not finding a significant difference in terms of leadership (top ranked citations) is quite reassuring for our potential sample selection: mobsters in the two samples appear to have the same visibility and importance in the organization.

We collected information about mafiosi and their closest neighbors, defined as all individuals recorded on the same page in the 1940 Census manuscript. Logan and Parman (2015) use this same strategy to identify neighbors in historical censuses, with the aim of tracking residential segregation over time. Census enumerators were assigned to one enumeration district (in most cases and all cases in larger cities) and they were provided with maps of their district and instructed to visit each house in order so that the Census cards that survive on ancestry.com show, in almost every instance, households directly adjacent to each other in their neighborhoods. Only 23 percent of the time in our sample were people recorded on the same page not living on the same exact street, but even in these cases they were still living very close by (in adjoining streets, for example, see Logan and Parman, 2015). A single Census card may contain more than one street, as the enumerator continued recording families "around the block". Because of the urban nature of mobsters, everybody that we identify as their neighbor is living at most one block away from them.

The advantage of the 1940 Census is that it allows for a search by first and last names as well as basic demographic characteristics and it was the first U.S. Census to ask questions at the individual level about highest grade of schooling attained, wage income, whether any non-labor income was earned in the previous year and migration in the past five years and it also provides information on the house value or rent paid for each household. The resulting database on each mobster provides information on their background and educational attainment during a period when they were already engaged in criminal activity (the only possible exceptions being mobsters in our database who were aged under 18 in 1940), as well as comparable measures of background for a group of their neighbors. We cleaned the data of typographic errors present on the ancestry.com Web site, to ensure a large enough sample size for the analyses below. The 1940 Census was only released to the public, with names, in 2012 and the FBN records were declassified and published in 2007, so this is the first time that such a dataset linking members of U.S. organized crime families and their illegal behaviors to earlier information on educational attainment and family background has been possible. Mastrobuoni (2015) provides more detail on the FBN source, but it contains information on approximately one quarter of mafia members in the 1960s.

While some of the neighbors might have been associated with the mafia, most were probably not. Of our 414 mobsters only in 5 instances did mobsters with different surnames share the same Census page with other mobsters: Joseph Filardo and Joseph Cusamano, Carlo Gambino and Gaetano Russo, Joseph Stracci and John Linardi, Agatino Garufi and Salvatore Maimone, Vincent Teriaca and Nicholas Bonina. In other words, only 10 out of 414 known mobsters lived close enough to end up on the same Census page.

Moreover, some of the differences in the characteristics of mobsters and neighbors suggest that neighbors are indeed less likely to be mobsters. We will see that neighbors are considerably less likely to be born in Italy (15 percent against 38 percent), they are also less likely to be employers or to be working on own account (12.95 percent against 31.14 percent) and twice as likely to be working for the government, probably as part of the New Deal Works Project Administration (9.92 percent against 4.72 percent). We will also see that they are less likely to underreport their income.

Here is how we selected mobsters and neighbors. Each mobster who, in that year, was not spending time in prison (32 out of 414 were incarcerated), was not attending school (45 out of 382 were still in school), and whose age was above 18 (6 out of 334 were minors) was then matched one to many with their white male neighbors selected as above and whose age is within 10 years of the mobster's age (we also use lower thresholds). The average number of records on each Census page (independently of race, gender and age) is 32.5 and never exceeds 40 because of the physical dimensions of the standard Census cards. In ninety percent of cases there are more than 25 such records. Once the sample is reduced as outlined above, the average number of neighbors is equal to 6.2.

This gives us a final sample of 311 mobsters and a comparable set of their peers, based on age, race, gender and place of residence, on which to run our analysis of the returns to education for criminals versus non-criminals. We firstly discuss summary characteristics of the samples before proceeding with our main analyses. We also acknowledge that we are running straightforward Mincer regressions using OLS. But, given that our data collection strategy allows us to observe mobsters and neighbors, in Section 5 we argue that the usually cited biases of such an approach should be present for both groups equally, allowing us to compare outcomes for the two. We additionally present estimates for broader comparison groups to investigate the validity of our mobster and neighbor findings.

4.2 Census Data

We draw various samples from the 1940 Federal Census IPUMS 1% sample, and investigate the returns to education in these samples for comparison with the estimated returns for mobsters. In line with the characteristics of mobsters, the most comprehensive sample includes all white men aged between 18 and 60. Then, we further select all U.S.-born men; all male citizens of the U.S.; all Italian-born men living in the U.S.; all immigrants living in the U.S.; second generation Italian men (defined as U.S.-born individuals having either a mother or father born in Italy). For robustness some specifications restrict further by age and weight the samples according to the age and location of mobsters across the U.S.

As for mobsters and neighbors, for all of these samples only men who are recorded as not attending school and being in the labor force are included. We then follow Goldin and Katz (2000) to multiply topcoded incomes by 1.5 and to construct our measure of potential experience used in the Mincer regressions below. This potential experience measure is equal to the minimum of (age-years of education- 7) and (age-15) and it thus reflects both the usual age of starting education during this period and the fact that historically a greater proportion of individuals reported implausibly low levels of educational attainment. The education variable for the 1940 Census, used to construct both our potential experience and education variables, records the highest grade of schooling attained.

In the analyses below we present results for all of these samples. It is not immediately obvious as to what the correct comparison group for high-level criminals should be, so we present results from the most general sample (all white men in the U.S.) down to a sample of white male neighbors using only those on the same Census card as each mobster. The most relevant comparisons are probably second-generation Italians, Italian immigrants and the neighbor sample. This is based on the demographics of the samples and the fact that, both historically and today, there is a high degree of residential clustering so that neighbors should share characteristics with mobsters and should face the same labor market and institutional constraints and opportunities.

4.3 Summary Statistics

In the analyses neighbors are always weighted by the inverse of their number $\omega_i = 1/n_i$, where the index i identifies mobsters. Table 2 shows the summary statistics for mobsters, for the matched neighbors, and for the entire sample of white men aged between 18 and 60 (the Appendix Tables A2 and A3 cover the remaining comparison groups).

Mobsters are considerably more likely to report no income (partly because they are more likely to be self-employed), and the ones they report are on average 20 percent lower than for matched neighbors. This is likely to be misreporting because it is incompatible with the value of the house where they live. Mobsters are more likely to own a house (33 percent vs. 32 percent), and their house is on average worth about 10 percent more compared to those of their neighbors and 65 percent more than the average white men of similar age and similar state of residence. Moreover, even mobsters who are renting tend to spend slightly more than their closest matches. Mobsters' neighbors, instead, are quite similar to the average white man, especially when the overall sample is weighted according to the distribution of age and state of residence of mobsters (and neighbors), the only exception being that they live in more expensive housing.

We note that the Census instructions were very clear that enumerators should explain that information was not allowed to be passed on to the authorities and they faced a penalty of up to a $1000 fine or 2 years imprisonment if they revealed any confidential information. Enumerators also swore to not accept answers that they believed to be false and to assert their authority in entering people's homes and finding the correct information. Each enumerator covered only 1-2 blocks and reported their work to a supervisor. It seems that these measures were successful enough for mobsters to be comfortable with being entered on a Census card. We still acknowledge that they are unlikely to have reported all of their income from all sources. In only 6 cases do we have 2 observations for a mobster's income– this comes from cases where we have the 1940 Census record and we were also able to find a prison admission register for the individual (but they were not in prison in 1940).[14] 2 of these cases were from the 1920s, but looking at those in the 1930s suggests that mobsters do underreport income. It is difficult, however,

[14]Admission registers which are searchable by name exist on ancestry.com for Sing Sing and Clinton prisons, both in New York state. They were consulted in December 2015, and can be found at the following urls: http://search.ancestry.co.uk/search/db.aspx?dbid=8922 and http://search.ancestry.co.uk/search/db.aspx?dbid=9023.

to tell if the 1940 income is reported much lower because of the time spent in jail or because of mobsters exercising more caution in responding to Census officials than prison officials.

Since underreporting might bias our results we will conduct our analysis considering three different measures of their economic status: i) income, ii) an index based on the value of their house and their monthly rent payments (called "housing value"), and iii) whether or not they own the house in which they live. Yet, the observed differences in education might be part of the story. Mobsters have, on average, one less year of education compared to their neighbors: 7.80 against 8.75. Their educational attainment is even lower compared to the average U.S. male aged 18-60 who had 9.05 years. In line with the anecdotal evidence outlined above, Figure 1 shows that this difference is mainly driven by differences in the likelihood of entering high school versus entering the labor market, when children are about 14.

In terms of other socioeconomic characteristics, mobsters are more likely to be foreigners (30 percent are aliens, while 25 percent have been naturalized compared to 22 percent and 18 percent respectively for the group of neighbors), and are more likely to have been born in Italy. They are more likely to be married but they have, on average, fewer children and live in smaller households. Geographic mobility is low for both groups: 85% of mobsters and 86% of their neighbors have lived in the same house for at least five years.

Figures 2 and 3, show the cumulative distributions of log income and log housing quality for mobsters and the main comparison group, neighbors. The housing quality measure is explained in more detail in the next section, but it is one measure of residential status constructed from the Census information on either rents paid (where an individual is a renter) or housing value (where an individual is a homeowner). The raw plots for mobsters and neighbors show that mobsters' reported log income is typically lower than the log income of neighbors. The opposite is true when looking at housing value. This is likely driven by income being underreported, an issue we return to in the next sections.

Now that we have introduced the data sources and some summary statistics for our samples of mobsters, neighbors, and others drawn from the 1940 Census, the following sections will outline the methodology and results, facilitating a comparison of the returns to education in legitimate and illegitimate activities.

5 Returns to Education

In order to establish the role that education plays in shaping earnings inside organized crime organizations, we follow the long tradition of Mincer-type regressions, estimated with OLS. The main drawback of this approach is the possibility of ability bias- that an omitted variable, such as unobserved ability, may be correlated with both log income and years of education. Our datasets do not allow us to utilize the IV approaches common in the literature, as outlined in the Literature Review section. However, we believe that the bias should be similar for the mobster sample and for the comparison samples of neighbors, Italians, second-generation Italians, U.S.-born and U.S.-citizen men. The existing literature has shown that OLS estimates are mostly biased downwards relative to those found using IV methods, so our results below may be considered lower bound estimates. We thus interpret our reported returns to education as reflecting both the true return to schooling and the component that reflects the fact that ability is correlated with years of schooling attained.

Below, we also address the problem of under-reporting of mobsters' income. We also present some specifications that control for the fact that we have a selected sample of criminals. By using the FBN data from 1960, we are starting with a sample of highly successful and well-connected mobsters and, since the FBN data allowed us to identify the connectedness and importance of each individual, we reweight some estimations to control for the fact that well-connected people are more likely to be in the sample, making it non-representative of the typical mobster.

5.1 Main Results

As is customary, we transform all outcomes into logarithms. There is an additional advantage of taking logarithms: if mobsters are only reporting a fraction of their income μY_i, in logarithms this fraction will be separated from the outcome $\log \mu + \log Y_i = \alpha + y_i$ and can be captured by the constant term (or by the other regressors). Taking logarithms we exclude all incomes that are zero. We will show that this is likely to improve the precision of the estimates as zero incomes do not appear to be genuinely zero but are rather driven by misreporting.

Next, we measure the returns to education for mobsters and neighbors. We follow the long tradition of Mincer regressions and use linear models, where the log of y is regressed on years of education and potential experience.[15] In order to allow errors to be correlated across mobsters residing in the same city standard errors are clustered at the city level.

The results for the (parsimonious) regressions of income when controlling only for potential experience and years of education are in Columns 1 and 5 of Table 3 for mobsters and our main comparison group, neighbors. Columns 2-4 and 6-8 introduce more controls (we label these our baseline specifications). Table 6 presents the returns to education for alternative comparison groups using both the parsimonious and baseline specifications. Later, we show that controlling for potential experience squared does not alter our main conclusions. We will also show that excluding the variables that could potentially be endogenous to income (marital status, number of household members, and number of children) does not alter the results.

[15] Heckman et al. (2003) use Census data for the period 1940-1990 to estimate flexible internal rates of return to schooling. They account for non-linearity in schooling, non-separability between schooling and work experience, etc. While they do find evidence of such non-linearities, the 1940 and 1950 Censuses provide support for Mincer's original, more basic, model.

The returns to education and potential experience are equal to 8.5 and 4.2 percent respectively, for mobsters, and 10.2 and 3.9 percent for neighbors. These are comparable to the 8-9% reported in the Clay et al. (2012) OLS model for data from the 1940 Census. It suggests that education has a similar payoff in illegitimate and legitimate activities when comparing groups living in close proximity and who face the same opportunities in legal and illegal work in 1940.[16]

For mobsters, the returns to education drop by 0.4 and 1.1 percentage points when adding other socioeconomic variables as well as city or state fixed effects. For their neighbors, adding socioeconomic variables reduces the returns by 0.9 percentage points, while city or state fixed effects have almost no effect on the estimates. Mobsters and neighbors have very similar returns to potential experience in all specifications.

As for the coefficients on the other controls, there are quite substantial differences between mobsters and neighbors. Mobsters born in Italy and those that have been naturalized tend to report much smaller incomes, while those that are immigrants (aliens in the Census) report higher incomes. For neighbors, none of these differences are significant.

To cope with the potential measurement error bias in income discussed above we employ two different strategies: we impute mobsters' income using neighbors' income and we estimate Mincer regressions using housing quality as the dependent variable.

Taking these strategies in turn, Table 4 displays results using a variety of imputation strategies to deal with missing and zero income data for mobsters. Imputation was used for all mobsters (the only exception is column 3 where we only impute zeros and missing data). The first column repeats the result from Table 3, Column 4, the city fixed effects specification. The second column imputes incomes using housing values and housing rents, which are arguably more difficult to misreport and are available for almost all mobsters.

[16]Given the low geographic mobility of individuals in the 1940 Census, demonstrated by the "location in 1935" variable, it is even likely that mobsters and neighbors may have attended the same schools. This will not be true for mobsters who were born in Italy and moved to the U.S. after their schooling was completed.

If y^n/h^n is the ratio between incomes and housing values for neighbors, we impute log income for mobsters using $log(\hat{y}^m) = log(y^n/h^n * h^m)$. Using mobsters' housing values the returns to education are almost unchanged, as shown in Column 2.

Starting with Column 3, we use neighbors' incomes as a proxy for mobsters' incomes. Not surprisingly, given the potential mismatch the returns to education tend to be smaller though still significant. In particular, in the "Only missing and zeros" column only the missing incomes are imputed ($log(\hat{y}^m) = log(y^n)$ when y^m is missing or zero) using neighbors' incomes weighted by the inverse sum of standardized absolute differences between mobsters and neighbors in the observable characteristics used in Table 3. In other words, more weight is given to neighbors that have similar characteristics to the mobsters.[17]

In the "All" column, the unweighted neighbors' income is used for all mobsters' income, and such fully imputed incomes are weighted in the "Weighting" column. In columns "Nationality" and "Nationality&Age", only neighbors of the same nationality (Italian or not), and whose age difference is smaller than 5 years, are used. These tighter matches deliver returns to education that are closer to the baseline results.

Turning to the second strategy, housing outcomes might be an alternative, possibly less distorted, proxy for long-term income or wealth. The housing outcomes for each household are almost never missing in the 1940 Census, for mobsters and others. Intuitively, since enumerators could see the asset in question, it's unlikely that responders would be able to under-report the value of housing. We compute *Housing Quality*, an index for housing quality that converts monthly rents into a house value. We used the ratio of house values to rents at the city level from the neighbor sample to convert mobster and neighbor rental

[17] More formally, if $d_{ik} = |x_{ik}^m - x_{ik}^n|$ is the distance between the mobsters' and the neighbors' observable characteristic $k = 1, ..., K$, then we define the standardized distance to be:

$$d_i = \sum_{k=1}^{K} \frac{d_{ik} - \bar{d}_k}{SD(d_k)}.$$

The weight for observation i, w_i, is simply equal to the inverse of the distance ($w_i = 1/d_i$).

values into housing quality proxies.[18] This index is then used as the dependent variable.

Table 5 shows that, even when using housing values, the returns to education and experience are quite comparable for the two groups. If anything, the returns seem to be larger for mobsters than for neighbors in the first two specifications (Columns 1 and 2). As in Table 3, for mobsters city fixed effects tend to reduce the coefficients on education (see Column 3). The reason may be that city fixed effects capture 57 percent of the variability of education for mobsters, as shown in the last row of the table. Mobsters appear to sort into cities based on their level of education more than their neighbors. In other words, a successful mobster tends to be based in big cities that offer more opportunities and have more expensive houses.[19]

We now place the mobster results in context by looking at other comparison groups beyond neighbors. We estimate returns for different groups: all men aged 18-60, men born in the U.S., US citizens, immigrants, men born in Italy, and second-generation Italian men. Although the counterfactuals that are plausibly the closest ones to our treatment group (mobsters) are that of men born in Italy and those with at least one parent born in Italy, we also use the other comparison groups to have other interesting benchmarks.

Table 6 shows the returns to education for the different comparison groups for the entire U.S., with Panel B weighting by mobsters' age and state of residence. Returns are highest for those born in the U.S. (11.8% in the weighted expanded specification) followed closely by U.S. citizens (10.4%). Second-generation Italians have an estimated return of 6.8%, immigrants 4.8% and lastly Italian immigrants 3.2%. These compare to 7.4% for mobsters using the same specification, putting them closest to the second-generation Italians. This group is likely a very close comparison in terms of both unobservable characteristics and market opportunities and so this result highlights that the returns to

[18]The rent ratio distribution was first censored at the top and bottom 5% marks.

[19]Table A1 in the Appendix shows returns to education when using home-ownership as the dependent variable. Education significantly affects the decision to buy a house only for neighbors and not for mobsters. The coefficient on education is always positive for mobsters but it is significant only in the specification with state fixed effects.

education in legal and illegal activities are very similar. Taken together, these results suggest that the low returns to investment in education for Italian immigrants in the legal labor market may partially explain both the low investment made by mobsters in education and their decision to become involved in organized crime.

That returns to education for mobsters are larger than for immigrants, and especially compared to the Italian immigrants (as shown by comparing results in Table 6 to those in Table 3) may not be surprising for a number of reasons. The mobster sample contains a variety of types of individuals– both immigrants who completed at least some of their education abroad and those born in the U.S. and individuals descended from Italian immigrants. This means that mobsters had more U.S.-specific education and training compared to the immigrant sample as a whole and that may partly explain the higher returns to their more relevant, U.S. education. It may also reflect the lack of opportunities in the legal labor market for Italian immigrants. Italians at the turn of the century were very concentrated in certain low-status occupations and had little access to union protections and this may have continued through 1940.

Finally, we re-estimate mobster returns to education, taking into account the importance of mobsters within the mafia organization. The FBN records contain information about the closest criminal associates. This allows us to build the network of mobsters. We deal with the incompleteness and non-randomness of such a network by modeling law enforcement's surveillance and detection of mafia mobsters as a Markov chain. More active, more important, and more connected mobsters are more likely to be noticed and tracked. This means that observed gangsters are likely to be more connected than the average one. If the only low-educated mobsters to come to the attention of authorities and entered into FBN files are those who happen to be quite successful, this would probably lead to downward biased estimates of the relationship between education and mobster outcomes.

In order to produce a representative sample of mobsters, "sampling" weights should underweight highly connected mobsters and vice versa. This kind of sampling resembles a procedure that is used to sample hidden populations, called snowball sampling (a detailed description of how to construct the sampling weights can be found in Mastrobuoni (2015)).

The weighting scheme gives less weight to individuals with high numbers of connections. This should control for selection into our mobster sample, and does not seem to affect the results regarding the returns to education. Column 1 of Table 7 presents results from specifications that are the same as Table 3, Column 4. Column 2 of that table reproduces the same specification but applies a Markov-Chain weighting on mobsters. Here, we assume that mobsters who were more senior and well-connected within the hierarchy of the mafia were those who were reported in the FBN files around 1960 and we want to control for this selection into our mobster sample. The next columns in the table show the effect of years of education on the likelihood of being self-employed (Columns 3 and 4 are estimated with the linear probability model with a dummy for being self-employed as the outcome). Then, Columns 5 and 6 report results using importance of a mobster within the organization as an outcome measure. The results show that education had a positive effect on both the likelihood of running a business in 1940 and of being more important/successful in the mafia organization in 1960, whether or not we use the Markov-Chain weighting.

In Table 8 we perform some robustness checks to be sure that our results are not biased by the particular specification we used. Column 1 show the estimates of our baseline regression. In Column 2 we use log hourly wage (a variable created dividing income by the total working hours per year, when the information is available) as our dependent variable. Returns to education are lower than in the baseline, but still significant. The results differ little when we control for a squared term of potential years of experience (Column 3) and when limiting the sample to neighbors whose age gap with mobsters is under 5 years (and not 10 as in the baseline) (see Column 4). We do this to explore the robustness of our results and to compare mobsters to an even more relevant comparison group. The results are also robust to focussing on individuals above age 22 (Column 5), an age at which individuals in the 1940s had usually completed their studies. We also estimate returns to education without using family controls (married, number of people in the households and number of children) that are potentially endogenous to education, but including all the other regressors as in the main specification (Columns 4 and 8 in Table 3). Again, we find results that are in line with our main specification. The gap of 1-2 percentage points between mobster and neighbor returns from Table 3 is mostly replicated in these robustness checks.

To make sure that our estimates do not differ too much when including people reporting zero income, the final robustness check uses a Poisson model. The returns to education for mobsters is still (weakly) significant with a lower coefficient (about 6 percent). This is consistent with the observed misreporting of zero income. Mobsters who report incomes of zero have on average higher levels of education, and live in more valuable housing. The opposite is true for neighbors.

In summary, the main message is that there was a substantial return to education for mafia members based on their information and activities in the 1940 Census. The magnitude of their returns is very similar to samples of their neighbors and to a more general sample of second-generation Italians. Immigrants, especially Italians, seem to have had very low returns to education in the legal labor market while those enjoying the highest returns were the U.S.-born and U.S. citizen groups. This is likely due in part to occupational clustering, with immigrants being found mostly at the bottom of the occupational ladder and finding it difficult to move upwards in the legal labor market.

5.2 Mechanisms

We discussed above that we believe the returns to education that we calculate for our mobster sample is so similar to the returns in legitimate activities found in the previous literature and in our various comparison samples presented in Tables 2 and 5 because mobsters, especially those who have been a member of the organization for a long time and are well connected, engage in crimes that require cognitive, organizational and even social skills of the type that are developed in formal education.

We are able to test this explanation more thoroughly, using information from the FBN files on the crimes with which individuals have been charged. In particular, we compare returns to education between business and non-business criminals, as defined above.

Our findings are reported in Table 9, with income as the outcome in Columns 1-4 and housing quality as the outcome in Columns 5-8. Even columns report returns to education and experience for these business criminals, while odd columns represent returns for all others. Returns to education are at least 9 percentage points higher for business criminals, and their returns to experience are at least 5 percentage points higher in the income regressions. These results are very consistent with our narrative that mobsters have surprisingly high returns to education because of the complex nature of the crimes and criminal network that they are involved in. Our main results can be considered an average of the results for these two types within the mobster group. This finding suggests that, in general, those engaged in complex crimes and white-collar crimes more generally may enjoy a high return to education. Conversely, the results for the non-business criminals within the mobster group reveal quite low returns to education, not very different to the immigrant and Italian immigrant returns to legitimate activities shown in Table 6. These are also closer to the low (but insignificant) returns reported by Carvalho and Soares (forthcoming) for their sample of young, lowly Rio gang members.

6 Conclusions

This paper analyzes the link between education and economic outcomes (income and housing values) for members of the Italian-American mafia, as well as for a sample of their neighbors and other comparison samples of white men living in the U.S. and located in the 1940 Census files. We focus on mobsters who were listed in a 1960 FBN publication and link these data with those of the 1940 Decennial Census.

Consistent with a career choice model, we find that the distribution of years of education of mobsters is first order stochastically dominated by the distribution of their neighbors. We also find that schooling has a positive return not only in legitimate activities, but also in illegitimate ones. In fact, mobsters' returns lie in between those for Italian and other immigrant groups (on the lower end) and those for U.S.-born individuals and U.S. citizens (on the higher end). While this might appear counterintuitive, a model of human capital investment where the working life is shorter (in this case because of expected prison time, injuries or death), predicts larger annual returns to education. This is because the marginal year of schooling will have to provide larger annual returns for somebody whose working life is shorter, as there will be less time to make up for the investment over the lifetime.[20]

The mafia business is usually a mix of legal and illegal activities. For illegal activities like racketeering, extortion, loan sharking, etc. skills acquired in education, like the ability to process numbers, to think logically, organize complicated logistics systems etc. are likely to increase with education and seem necessary for success in these mafia roles. Moreover, often times loan sharking would allow mobsters to acquire legitimate businesses convenient for money-laundering purposes. Returns to education in such activities are likely to be large too. The FBN records allowed us to identify the types of crimes mobsters were charged with and investigate returns to education and experience for "business"

[20]This rests on the assumption that education either signals productivity or contributes to the productivity of mobsters.

criminals, who were involved in the types of activities just described, and others, who were not charged with such crimes. The results are fully consistent with our narrative and show returns to education for business criminals that are of the same order of magnitude as for the U.S.-born and U.S. citizens.

We conclude that, at least for career criminals operating at a high level in complex organizations who perpetrate serious crimes, education is quite valuable.

This study has focused on a very specific organized crime group, the mafia. Whether these results hold up in other criminal organizations, with more or less complex business models, is a possible avenue of future research. Similarly, this paper is silent on the value of education for criminals at lower ranks within a larger criminal organization and, given the small sample size and unstable results in Carvalho and Soares (forthcoming), the returns for such individuals remain uncertain.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2014) 'A nation of immigrants: Assimilation and economic outcomes in the age of mass migration.' *Journal of Political Economy* 122(3), 467–506

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

Anderson, Mark D. (2014) 'In school and out of trouble? The minimum dropout age and juvenile crime.' *Review of Economics and Statistics* 96(2), 318–331

Ashenfelter, Orley, Colm Harmon, and Hessel Oosterbeek (1999) 'A review of estimates of the schooling/earnings relationship, with tests for publication bias.' *Labour economics* 6(4), 453–470

Bowles, Samuel, and Herbert Gintis (2002) 'Schooling in capitalist america revisited.' *Sociology of education* pp. 1–18

Buonanno, Paolo, Ruben Durante, Giovanni Prarolo, and Paolo Vanin (2015) 'Poor Institutions, Rich Mines: Resource Curse and the Origins of the Sicilian Mafia.' *The Economic Journal* 125(586), F175–F202

Card, David (1999) 'The causal effect of education on earnings.' *Handbook of labor economics* 3, 1801–1863

_ (2001) 'Estimating the return to schooling: Progress on some persistent econometric problems.' *Econometrica* 69(5), 1127–1160

Carvalho, Leandro, and Rodrigo Soares (forthcoming) 'Living on the edge: Youth entry, career and exit in drug-selling gangs.' *Journal of Economic Behavior & Organization*

Clay, Karen, Jeff Lingwall, and Melvin Stephens Jr (2012) 'Do schooling laws matter? evidence from the introduction of compulsory attendance laws in the united states.' Technical Report, National Bureau of Economic Research

Collins, William J, and Marianne H Wanamaker (2013) 'Selection and economic gains in the great migration of african americans: New evidence from linked census data.' Technical Report, National Bureau of Economic Research

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

Ehrlich, Isaac (1975) 'On the relation between education and crime.' In 'Education, Income, and Human Behavior' (NBER) pp. 313–338

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

Ferrante, Louis (2011) *Mob Rules: What The Mafia Can Teach The Legitimate Businessman* (Penguin Group)

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

Goldin, Claudia, and Lawrence F Katz (2000) 'Education and income in the early twentieth century: Evidence from the prairies.' *The Journal of Economic History* 60(03), 782–818

_ (2008a) 'Mass secondary schooling and the state: The role of state compulsion in the high school movement.' *NBER Chapters* pp. 275–310

Goldin, Claudia, and Lawrence Katz (2008b) *The race between education and technology* (Cambridge, MA: Belknap Press of Harvard University Press)

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

Heckman, James J, Lance J Lochner, and Petra E Todd (2003) 'Fifty years of mincer earnings regressions.' Technical Report, National Bureau of Economic Research

Hjalmarsson, Randi, Helena Holmlund, and Matthew J Lindquist (2015) 'The effect of education on criminal convictions and incarceration: Causal evidence from micro-data.' *The Economic Journal*

ISTAT (2014) 'Censimenti e Societá, Mutamenti Sociodemografici della Sicilia in 150 Anni di Storia.' Temi e Statistiche, Istituto Nazionale di Statistica

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

Lochner, Lance (2004) 'Education, work, and crime: A human capital approach.' *International Economic Review* 45(3), 811–843

_ (2011) 'Nonproduction benefits of education: Crime, health, and good citizenship.' *Handbook of the Economics of Education* 4, 183

Lochner, Lance, and Enrico Moretti (2004) 'The effect of education on crime: Evidence from prison inmates, arrests, and self-reports.' *American Economic Review* 94(1), 155–189

Logan, Trevon, and John Parman (2015) 'The national rise in residential segregation.' Technical Report, National Bureau of Economic Research

Luallen, Jeremy (2006) 'School's out... forever: A study of juvenile crime, at-risk youths and teacher strikes.' *Journal of urban economics* 59(1), 75–103

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

Maas, Peter (1968) *The Valachi Papers* (Putnam, New York)

Machin, Stephen, Olivier Marie, and Sunčica Vujić (2011) 'The crime reducing effect of education.' *The Economic Journal* 121(552), 463–484

Mastrobuoni, Giovanni (2015) 'The Value of Connections: Evidence from the Italian-American Mafia.' *The Economic Journal* 125(586), F256–F288

Meghir, Costas, Mårten Palme, and Marieke Schnabel (2012) 'The effect of education policy on crime: an intergenerational perspective.' Technical Report, National Bureau of Economic Research

Mincer, Jacob (1974) 'Schooling, experience, and earnings. human behavior & social institutions no. 2.'

Pretelli, Matteo (2013) 'Italians and Italian Americans, 1870-1940.' In 'Immigrants in American History: Arrival, Adaptation, and Integration' (Abc-Clio) pp. 437–449

Roemer, William F (1995) *Accardo: The genuine godfather* (Donald I. Fine)

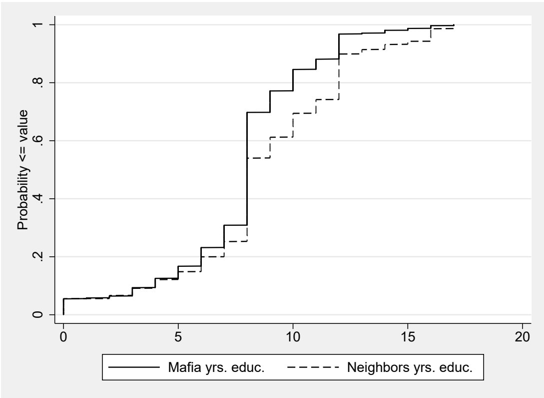

Figure 1: Cumulative Distribution of Years of Education

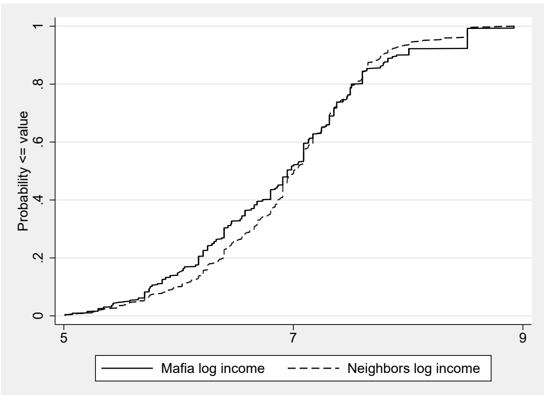

Figure 2: Cumulative Distribution of (log) Income

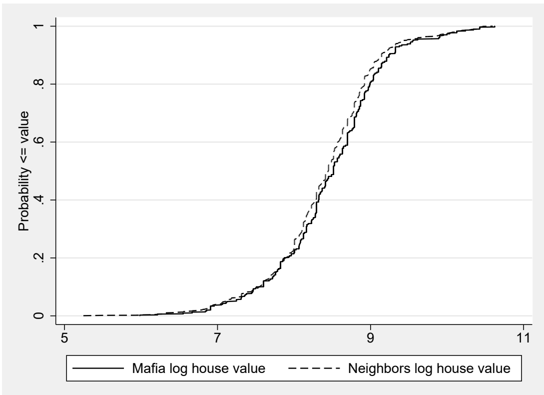

Figure 3: Cumulative Distribution of (log) Housing Quality

Table 1: Summary statistics for mobsters

	Unmatched		Matched			
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Diff.	t-stat
Year of birth	1912.17	8.15	1910.23	6.93	1.94	-3.92
Business criminal	0.23	0.42	0.31	0.46	0.07	2.30
Top ranked citations	0.83	1.01	0.87	1.05	-0.04	0.61
Married	0.71	0.45	0.80	0.40	-0.08	2.19
Divorced	0.07	0.25	0.07	0.26	0.00	0.11
# of children	0.90	1.32	1.17	1.58	-0.27	2.32
No arrests	0.17	0.37	0.15	0.36	0.02	-0.75
First arrest	1932.99	10.88	1931.94	10.61	1.05	-1.21
Residing in NY	0.41	0.49	0.48	0.50	-0.06	1.47
Residing in NJ	0.05	0.22	0.07	0.25	-0.01	0.78

Notes: We compare the characteristics from the FBN records of mobsters that were matched with the 1940 CENSUS and were selected ($nobs = 330$) with those that were not matched, or were matched but could not be selected ($nobs = 471$). The last column tests the difference between the two means.

Table 2: Summary statistics

	Mafia Mobsters			Neighbors			White men aged 18 to 60				
							Unweighted		Weighted		
	Mean	St. Dev.	N	Mean	St. Dev.	N	Mean	St. Dev.	Mean	St. Dev.	N
Income	825	1,203	291	1,070	1,184	1,782	858	954	1,081	1,006	337,021
Zero income	0.38	0.49	291	0.23	0.42	1,782	0.27	0.45	0.19	0.40	337,021
Home owner	0.33	0.47	311	0.32	0.47	1,892	0.43	0.50	0.30	0.46	337,021
House value	7,336	10,147	104	6,651	10,311	533	3,151	3,233	4,383	3,770	146,473
Rent	36.37	26.14	200	35.5	29.1	1,267	25.14	27.83	35.35	26.86	178,824
Housing	6,622	7,331	304	6,210	7,259	1,800	3329	3576	4,756	3,765	325,297
Yrs. of education	7.80	3.08	311	8.75	3.71	1,892	9.05	3.43	9.69	3.29	337,021
Age in years	33.86	8.45	311	34.18	9.75	1,892	36.79	11.65	33.07	8.10	337,021
Married	0.73	0.45	311	0.64	0.48	1,892	0.68	0.46	0.65	0.48	337,021
Born in Italy	0.38	0.49	311	0.15	0.36	1,892	0.02	0.15	0.04	0.20	337,021
Alien citizen	0.30	0.46	311	0.22	0.41	1,892	0.12	0.32	0.18	0.38	337,021
Naturalized citizen	0.25	0.43	311	0.18	0.39	1,892	0.09	0.28	0.13	0.34	337,021
Self employed	0.33	0.47	311	0.03	0.17	1,892	0.09	0.28	0.20	0.40	337,021
# of HH members	2.88	2.15	311	3.42	2.02	1,892	4.08	2.32	3.86	2.10	337,021
# of children	0.51	1.03	311	0.76	1.21	1,892	1.20	1.70	0.92	1.38	337,021
Same residence last 5yrs.	0.85	0.36	311	0.86	0.35	1,892	0.80	0.40	0.84	0.37	337,021

Notes: The Neighbors as well as all comparison groups are restricted to white men aged between 18 and 60. Reweighting is done according to the distribution of mobsters by age and state of residence.

Table 3: Mincer Regressions Using Yearly Income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	log-Income of Mafia members				log-Income of Neighbors			
Yrs. of education	0.085***	0.081***	0.073***	0.074***	0.102***	0.093***	0.091***	0.090***
	(0.016)	(0.017)	(0.016)	(0.020)	(0.008)	(0.008)	(0.009)	(0.008)
Potential yrs. of experience	0.042***	0.039***	0.033***	0.028***	0.039***	0.028***	0.028***	0.028***
	(0.008)	(0.013)	(0.009)	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)
Married		0.232	0.181	0.247		0.478***	0.462***	0.453***
		(0.182)	(0.181)	(0.194)		(0.074)	(0.067)	(0.048)
Born in Italy		-0.432	-0.573**	-0.835***		-0.015	-0.032	-0.020
		(0.271)	(0.234)	(0.159)		(0.084)	(0.077)	(0.073)
Alien citizen		0.916**	1.150***	1.365***		-0.044	-0.071	-0.029
		(0.376)	(0.331)	(0.228)		(0.078)	(0.077)	(0.088)
Naturalized citizen		-0.662**	-0.764***	-0.771***		0.018	0.042	0.002
		(0.274)	(0.277)	(0.247)		(0.071)	(0.076)	(0.078)
# of HH members		-0.085	-0.139**	-0.187***		-0.012	-0.025	-0.031
		(0.075)	(0.055)	(0.029)		(0.026)	(0.021)	(0.020)
# of children		0.150	0.226*	0.348***		-0.027	-0.009	0.001
		(0.153)	(0.115)	(0.052)		(0.029)	(0.025)	(0.022)
Same residence last 5 yrs.		-0.047	0.024	0.079		0.070	0.057	0.017
		(0.180)	(0.143)	(0.088)		(0.115)	(0.118)	(0.126)
Constant	5.363***	5.525***	5.779***	5.867***	4.844***	4.898***	4.976***	5.027***
	(0.444)	(0.630)	(0.613)	(0.726)	(0.161)	(0.205)	(0.193)	(0.208)
State fixed effects			Y				Y	
City fixed effects				Y				Y
Observations	180	180	180	180	1381	1381	1381	1381
R-squared	0.083	0.131	0.256	0.466	0.206	0.265	0.298	0.383

Notes: There are a total of 311 mobsters in the data. The number of observations refer to the unweighted data. There are a total of 180 mobsters with positive incomes. Mincer wage regressions with clustered (by city) standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: Mincer Regressions Using Imputed Income

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			log-Income				
Imputation:	None	From housing	Only missing and zeros	All	Weighting	Nationality	Nationality&Age
Yrs. of education	0.074***	0.075***	0.063***	0.043***	0.046***	0.061***	0.068***
	(0.020)	(0.014)	(0.009)	(0.012)	(0.013)	(0.013)	(0.010)
Potential years of experience	0.028***	0.045***	0.031***	0.039***	0.041***	0.058***	0.061***
	(0.007)	(0.006)	(0.007)	(0.008)	(0.008)	(0.010)	(0.010)
Imputed zero incomes			0.354**				
			(0.172)				
Other Xs	Y	Y	Y	Y	Y	Y	Y
City fixed effects	Y	Y	Y	Y	Y	Y	Y
Observations	180	1,304	1,631	1,380	1,380	902	569
R-squared	0.466	0.213	0.370	0.284	0.273	0.296	0.345

Notes: This table shows returns to education using seven different imputation strategies for income (one for each column of the table). Imputation was used for all mobsters (the only exception is column 3 where we only impute zeros and missing data).
"None": this is our baseline regression with city fixed effects (see Column 4 in Table 3).
"From housing": we use the ratio between income and housing value (and rent) for neighbors at city level (y^n/h^n) to impute income for all mobsters ($log(\hat{y}^m) = log(y^n/h^n * h^m)$).
"Only missing and zeros": missing data on income as well as the zeros are imputed using neighbors' incomes weighted by the inverse sum of standardized absolute differences between mobsters and neighbors in the observable characteristics used in Table 3. More weight is given to neighbors that have similar characteristics to the mobsters (see footnote 16 for more details).
"All": the unweighted neighbors' income is used for all mobsters' income (not just for the missing values and the zeros).
"Weighting": the imputed incomes in the "All" column are weighted using the same weights as in Column 3.
"Nationality": only neighbors of the same nationality (Italian or not) are used to impute mobsters' income.
"Nationality&Age": only neighbors of the same nationality and whose age difference is smaller than 5 years are used to impute mobsters' income. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Mincer Regressions Using Housing Quality

	(1)	(2)	(3)	(4)	(5)	(6)
	log-Housing of Mafia members			log-Housing of Neighbors		
Yrs. of education	0.064***	0.059***	0.034**	0.057***	0.057***	0.043***
	(0.015)	(0.016)	(0.015)	(0.009)	(0.007)	(0.008)
Potential yrs. of experience	0.021***	0.022***	0.015**	0.014***	0.016***	0.012***
	(0.006)	(0.005)	(0.006)	(0.004)	(0.003)	(0.004)
Married		0.010	0.094		0.109**	0.128***
		(0.075)	(0.069)		(0.043)	(0.042)
Born in Italy		-0.213**	-0.235*		-0.156*	-0.216**
		(0.100)	(0.126)		(0.091)	(0.100)
Alien citizen		0.207	0.182		0.083	0.108
		(0.153)	(0.216)		(0.100)	(0.084)
Naturalized citizen		-0.183	-0.139		-0.051	-0.043
		(0.154)	(0.210)		(0.095)	(0.072)
# of HH members		-0.046*	-0.041		0.016	0.019
		(0.027)	(0.029)		(0.016)	(0.014)
# of children		0.067*	0.089**		-0.039	-0.033
		(0.040)	(0.039)		(0.036)	(0.034)
Same residence last 5 yrs.		-0.007	0.387**		-0.097	-0.099
		(0.194)	(0.178)		(0.083)	(0.060)
Constant	7.599***	7.781***	7.684***	7.642***	7.624***	7.777***
	(0.295)	(0.210)	(0.281)	(0.221)	(0.135)	(0.159)
State fixed effects		Y			Y	
City fixed effects			Y			Y
Observations	304	304	304	1,800	1,800	1,800
R-squared	0.059	0.263	0.505	0.054	0.234	0.413
Explained variability of education			0.570			0.375

Notes: There are a total of 311 mobsters in the data. The number of observations refer to the unweighted data. Weighting, there are 311 observations in both groups. Mincer wage regressions with clustered (by city) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Returns to Education for Various Subsamples

	Sample	Average		Model	Returns to education	
		Income	Years of educ.		Income	Housing
		Panel A: Unweighted				
(1)	All aged 18-60	857.7	9.053	Parsimonious	0.116***	0.121***
(2)				Baseline	0.105***	0.108***
(3)	Born in the US	847	9.348	Parsimonious	0.131***	0.144***
(4)				Baseline	0.116***	0.120***
(5)	US citizen	860.5	9.168	Parsimonious	0.120***	0.127***
(6)				Baseline	0.108***	0.111***
(7)	Immigrants	927.6	7.105	Parsimonious	0.057***	0.058***
(8)				Baseline	0.049***	0.051***
(9)	Italians	785.4	5.321	Parsimonious	0.030***	0.034***
(10)				Baseline	0.029***	0.033***
(11)	Second generation Italians	776.4	9.256	Parsimonious	0.069***	0.059***
(12)				Baseline	0.064***	0.062***
	Panel B: Weighted by mobsters' age and State of residence					
(13)	All aged 18-60	1081	9.691	Parsimonious	0.098***	0.073***
(14)				Baseline	0.089***	0.078***
(15)	Born in the US	1086	10.11	Parsimonious	0.118***	0.093***
(16)				Baseline	0.106***	0.094***
(17)	US citizen	1090	9.822	Parsimonious	0.104***	0.078***
(18)				Baseline	0.093***	0.082***
(19)	Immigrants	1059	8.022	Parsimonious	0.048***	0.044***
(20)				Baseline	0.040***	0.040***
(21)	Italians	878	6.254	Parsimonious	0.032***	0.035***
(22)				Baseline	0.031***	0.035***
(23)	Second generation Italians	850.3	9.152	Parsimonious	0.068***	0.049***
(24)				Baseline	0.061***	0.052***

Notes: Each row corresponds to two separate linear regressions, using either income or housing quality as dependent variables. The parsimonious regressions control for just potential years of experience while the baseline regressions adds city fixed effects as well as all the socioeconomic controls (as in Column 4 of Table 3). The weighted regressions use the mobsters' distribution by age and states to reweight the data. Mincer wage regressions with clustered (by city) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Mincer Regressions for Additional Outcomes, Weighting for Sampling Design

	(1)	(2)	(3)	(4)	(5)	(6)
	Log-income		Self-employed		Top ranked citations	
Years of education	0.074***	0.078***	0.018**	0.030***	0.058**	0.037*
	(0.020)	(0.027)	(0.009)	(0.008)	(0.024)	(0.019)
Potential yrs. of experience	0.028***	0.030***	0.014**	0.017**	0.028***	0.035***
	(0.007)	(0.007)	(0.006)	(0.007)	(0.008)	(0.009)
Married	0.247	0.372	0.027	0.015	-0.000	0.193
	(0.194)	(0.371)	(0.081)	(0.052)	(0.147)	(0.149)
Born in Italy	-0.835***	-0.378	0.024	0.101	-0.159	-0.294
	(0.159)	(0.297)	(0.121)	(0.124)	(0.195)	(0.211)
Alien citizen	1.365***	0.508	0.158	0.195	0.604*	0.095
	(0.228)	(0.486)	(0.221)	(0.264)	(0.321)	(0.471)
Naturalized citizen	-0.771***	-0.337	-0.119	-0.315	-0.562*	-0.077
	(0.247)	(0.503)	(0.161)	(0.203)	(0.302)	(0.417)
# of HH members	-0.187***	-0.113**	-0.013	-0.005	-0.023	0.045
	(0.029)	(0.043)	(0.011)	(0.020)	(0.044)	(0.052)
# of children	0.348***	0.352***	0.008	-0.010	0.069	-0.077
	(0.052)	(0.099)	(0.027)	(0.045)	(0.061)	(0.087)
Same residence last 5 yrs.	0.079	-0.095	0.131	0.080	0.220	0.241
	(0.088)	(0.198)	(0.102)	(0.074)	(0.143)	(0.148)
Constant	5.867***	5.723***	-0.210	-0.338*	-0.156	-0.458
	(0.726)	(0.399)	(0.146)	(0.193)	(0.343)	(0.389)
Markov Chain weighting	No	Yes	No	Yes	No	Yes
Observations	181	181	344	344	344	344
R-squared	0.466	0.431	0.469	0.551	0.331	0.375
Average dependent variable	6.758	6.762	0.332	0.323	0.908	0.754

Notes: All Mincer-style regressions are estimated using linear models, and control for city fixed effects as well as for all the controls listed in Table 3. The variable "Top ranked citations" counts the number of words in the records that describe a mobster as a top ranked member of the Mafia, while self-employed is a dummy that is equal to one if individuals declare to be self-employed in the 1940 Census. The Markov-chain weighting gives more weight to mobsters who in the FBN records appear to have few connections. Clustered (by city) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Robustness Mincer Wage Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				log-Income			
Panel A: Mafia Members	*Baseline*	*log(Income/Tot.Hrs)*	*w. Exp. sq.*	$\Delta age \leq 5$	$age \geq 22$	*No family controls*	*Poisson*
Mafia yrs. educ.	0.081***	0.067***	0.082***	0.100***	0.067***	0.078***	0.063**
	(0.017)	(0.018)	(0.017)	(0.026)	(0.018)	(0.018)	(0.025)
Potential yrs. of experience	0.039***	0.022	0.041***	0.055***	0.035**	0.046***	0.036**
	(0.013)	(0.015)	(0.014)	(0.017)	(0.016)	(0.009)	(0.017)
Pot. yrs. of experience squared			-0.001**				
			(0.000)				
Observations	180	128	180	99	166	180	291
R-squared	0.131	0.124	0.137	0.206	0.097	0.418	
log-likelihood							-189382
Panel B: Neighbors	*Baseline*	*log(Income/Tot.Hrs)*	*w. Exp. sq.*	$\Delta age \leq 5$	$age \geq 22$	*No family controls*	*Poisson*
Neighbors yrs. educ.	0.093***	0.067***	0.095***	0.097***	0.088***	0.101***	0.088***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
Potential yrs. of experience	0.028***	0.020***	0.034***	0.036***	0.021***	0.039***	0.023***
	(0.004)	(0.003)	(0.004)	(0.006)	(0.003)	(0.004)	(0.004)
Pot. yrs. of experience squared			-0.001***				
			(0.000)				
Observations	1380	1123	1380	814	1278	1380	1782
R-squared	0.256	0.160	0.277	0.262	0.193	0.199	.
log-likelihood							-141845

Notes: There are a total of 311 mobsters in the data. The number of observations refer to the unweighted data. Weighting, there are 311 observations in both groups. Mincer wage regressions with clustered (by city) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Mincer Wage Regressions by Type of Criminal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		log-Income				log-Housing Quality		
"Business" criminals	No	Yes	No	Yes	No	Yes	No	Yes
Years of education	0.063***	0.112	0.056***	0.161*	0.050**	0.141**	0.052***	0.140**
	(0.021)	(0.070)	(0.019)	(0.089)	(0.020)	(0.058)	(0.018)	(0.067)
Potential yrs. of experience	0.026***	0.079***	0.025**	0.117***	0.022**	0.044**	0.030***	0.032**
	(0.009)	(0.013)	(0.011)	(0.016)	(0.010)	(0.018)	(0.010)	(0.015)
Other Xs and State fixed effects	No		Yes		No		Yes	
Observations	132	50	128	50	153	58	149	58
R-squared	0.033	0.358	0.232	0.669	0.059	0.149	0.419	0.338

Notes: "Business" criminals have committed at least one of the following crimes: embezzlement, forgery, fraud, counterfeiting, gambling, prostitution, tax evasion, and bookmaking. "Non business" crimes include, among others: robbery, murder, weapons offenses, simple assault, larceny, burglary, drug and liquor offenses. Mincer wage regressions with clustered (by city) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

A Appendix

Table A1: Mincer Regressions Using Home-Ownership

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Homeowner (Mafia)				Homeowner (Neighbors)			
Yrs. of education	0.043	0.044	0.075**	0.006	0.046***	0.056***	0.053***	0.048***
	(0.028)	(0.028)	(0.029)	(0.024)	(0.011)	(0.011)	(0.010)	(0.009)
Potential yrs. of experience	0.031***	0.063***	0.073***	0.065***	0.018***	0.030***	0.025***	0.024**
	(0.008)	(0.009)	(0.010)	(0.011)	(0.007)	(0.008)	(0.007)	(0.010)
Married		-0.843***	-1.046***	-1.039***		-0.259***	-0.264***	-0.342***
		(0.218)	(0.250)	(0.342)		(0.089)	(0.090)	(0.093)
Born in Italy		-0.536**	-0.605**	-0.450		0.240	0.305*	0.352*
		(0.227)	(0.277)	(0.295)		(0.156)	(0.160)	(0.207)
Alien citizen		0.381	0.412	0.039		-0.595***	-0.610***	-0.574***
		(0.439)	(0.569)	(0.789)		(0.207)	(0.217)	(0.202)
Naturalized citizen		-0.066	-0.189	0.138		0.332**	0.464***	0.437***
		(0.389)	(0.505)	(0.734)		(0.160)	(0.178)	(0.143)
# of HH members		0.022	0.005	-0.011		0.091***	0.114***	0.111***
		(0.043)	(0.042)	(0.046)		(0.035)	(0.034)	(0.036)
# of children		0.023	0.053	0.019		-0.073	-0.115**	-0.129**
		(0.078)	(0.079)	(0.069)		(0.058)	(0.053)	(0.059)
Same residence last 5 yrs.		0.180	0.241	0.583		-0.046	0.088	0.382**
		(0.241)	(0.238)	(0.556)		(0.123)	(0.129)	(0.149)
Constant	-1.322***	-1.446***	-1.660***	-5.259***	-1.216***	-1.538***	-1.154***	-5.851***
	(0.306)	(0.338)	(0.471)	(0.450)	(0.326)	(0.263)	(0.247)	(0.347)
State fixed effects			Y				Y	
City fixed effects				Y				Y
Observations	311	311	308	237	1,892	1,892	1,892	1,822
log-likelihood	-192.9	-179.8	-160.9	-115.5	-190.5	-186.1	-173.5	-146.9

Notes: There are a total of 311 mobsters in the data. The number of observations refer to the unweighted data. Weighting, there are 311 observations in both groups. Mincer wage regressions with clustered (by city) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Summary Statistics for US Men and for US Citizens

	US men					Citizens				
	Unweighted		Weighted			Unweighted		Weighted		
	Mean	St. Dev.	Mean	St. Dev.	N	Mean	St. Dev.	Mean	St. Dev.	N
Income	847	957	1,086	1,015	292,629	861	959	1,090	1,012	323,300
Zero income	0.27	0.45	0.19	0.39	292,629	0.27	0.45	0.19	0.39	323,300
Home owner	0.44	0.50	0.33	0.47	292,629	0.44	0.50	0.31	0.46	323,300
House value	3,084	3,226	4,311	3,726	128,352	3,151	3,234	4,374	3,739	141,763
Rent	24.25	27.92	34.82	27.45	154,548	25.06	27.93	35.34	27.00	170,695
Housing	3,230	3,574	4,678	3,809	282,900	3,321	3,581	4,750	3,770	312,458
Yrs. of education	9.35	3.24	10.11	2.98	292,629	9.17	3.34	9.82	3.19	323,300
Age in years	35.62	11.47	31.96	7.84	292,629	36.50	11.62	32.84	8.07	323,300
Married	0.67	0.47	0.62	0.49	292,629	0.68	0.47	0.65	0.48	323,300
Born in Italy	0.00	0.00	0.00	0.00	292,629	0.02	0.13	0.03	0.17	323,300
Alien citizen	0.00	0.00	0.00	0.00	292,629	0.09	0.28	0.14	0.35	323,300
Naturalized citizen	0.00	0.00	0.00	0.00	292,629	0.09	0.28	0.14	0.35	323,300
# of HH members	4.10	2.31	3.91	2.12	292,629	4.09	2.30	3.89	2.09	323,300
# of children	1.13	1.65	0.84	1.32	292,629	1.18	1.68	0.91	1.36	323,300
Same residence last 5yrs.	0.79	0.41	0.83	0.37	292,629	0.80	0.40	0.84	0.36	323,300

Notes: All comparison groups are restricted to white men aged between 18 and 60. Reweighting is done according to the distribution of mobsters by age and state of residence.

Table A3: Summary Statistics for Immigrants, Italians, and Second-Generation Italian Men

	Immigrants					Italians					Second generation Italians				
	Unweighted		Weighted			Unweighted		Weighted			Unweighted		Weighted		
	Mean	St. Dev.	Mean	St. Dev.	N	Mean	St. Dev.	Mean	St. Dev.	N	Mean	St. Dev.	Mean	St. Dev.	N
Income	928	926	1,059	970	44,374	776	717	878	781	7,694	928	926	850	755	2,260
Zero income	0.26	0.44	0.22	0.41	44,374	0.22	0.42	0.23	0.42	7,694	0.26	0.44	0.20	0.40	2,260
Home owner	0.41	0.49	0.22	0.41	44,374	0.32	0.47	0.27	0.45	7,694	0.41	0.49	0.21	0.41	2,260
House value	3,623	3,240	4,810	3,997	18,114	3,757	2,986	4,347	3,188	3,397	3,623	3,240	4,764	3,527	730
Rent	30.80	26.56	37.19	24.64	24,267	26.91	20.27	31.04	23.73	4,128	30.80	26.56	29.66	17.28	1,372
Housing	3,990	3,517	5,072	3,561	42,381	3,803	2,879	4,331	3,262	7,525	3,990	3,517	4,268	2,706	2,102
Yrs. of education	7.11	3.96	8.02	3.89	44,374	9.26	2.72	6.25	3.61	7,694	7.11	3.96	9.15	2.68	2,260
Age in years	44.46	9.79	37.49	7.59	44,374	29.22	8.24	38.13	7.82	7,694	44.46	9.79	29.10	6.27	2,260
Married	0.78	0.41	0.75	0.43	44,374	0.54	0.50	0.81	0.39	7,694	0.78	0.41	0.60	0.49	2,260
Born in Italy	0.17	0.38	0.20	0.40	44,374	0.00	0.00	1.00	0.00	7,694	0.17	0.38	0.00	0.00	2,260
Alien citizen	0.89	0.31	0.90	0.29	44,374	0.00	0.00	0.91	0.28	7,694	0.89	0.31	0.00	0.00	2,260
Naturalized citizen	0.65	0.48	0.66	0.48	44,374	0.00	0.00	0.70	0.46	7,694	0.65	0.48	0.00	0.00	2,260
# of HH members	3.97	2.36	3.68	2.00	44,374	3.81	2.26	4.40	2.20	7,694	3.97	2.36	3.64	2.07	2,260
# of children	1.67	1.95	1.25	1.55	44,374	0.51	0.97	1.83	1.90	7,694	1.67	1.95	0.57	1.02	2,260
Same residence last 5yrs.	0.86	0.35	0.86	0.34	44,374	0.87	0.34	0.92	0.27	7,694	0.86	0.35	0.91	0.29	2,260

Notes: All comparison groups are restricted to white men aged between 18 and 60. Reweighting is done according to the distribution of mobsters by age and state of residence.