ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 763 February 2015

Returns to Education and Experience in Criminal Organizations: Evidence from the Italian-American Mafia

Nadia Campaniello, Rowena Gray and Giovanni Mastrobouni

Returns to Education and Experience in Criminal Organizations: Evidence from the Italian-American Mafia*

Nadia Campaniello[†], Rowena Gray[‡], Giovanni Mastrobuoni[§]

February 2015[¶]

Abstract

Is there any return to education in criminal activities? This is the first paper that investigates 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 have been capable businessmen, orchestrating crimes that require 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 organising supply chains, logistics and distribution when setting up a drug dealing system. We address this question by comparing mobsters with their closest (non-mobster) neighbors using United States Census data in 1940. We document that mobsters have one year less education than their neighbors on average. None of the specifications presented identified any significant difference in the returns to education between these two groups. Private returns to education exist also in the illegal activities characterised by a certain degree of complexity as in the case of organized crime in mid-twentieth century United States.

Keywords: Returns to education; organized crime
JEL classification codes: I26; N32; K42; Z13

*We acknowledge the excellent research assistance of Patryk Bronka and Gautham Udupa.

[†]Department of Economics, University of Essex, 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 Netspar, Email: gmastrob@essex.ac.uk.

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 723 mobsters belonging to the Italian-American mafia with the 1940 Census. This gives us information about income, housing values, education, job characteristics, as well as the precise address of residence. We match these mobsters with their white, male neighbors of similar age who, we argue, are unlikely to be mafia associates.

Economists have shown that increased levels of education reduce criminal participation. This implies that education is likely to be more valued 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 their neighbors.

But this finding does not imply that 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* returns to education.

This is true unless the extra schooling is not marketable. So, is schooling marketable in the mafia? This question is really about the mafia's complex business model and about 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 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 (setting up supply chains), etc. It is ultimately an empirical question 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.

We estimate Mincer-type regressions using log income, log housing value, and log rent as outcomes. The main variable of interest is years of education.

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. In Section 5 we discuss that the biases that are usually associated with returns to education, ability bias and measurement error bias, would most likely reduce such returns compared to the returns of the mobsters' neighbors.

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.

The paper proceeds as follows. We first discuss the existing literature on education and crime, before providing an overview of the history of mafia organizations and members in the United States before 1960. We then present our novel dataset and empirical methodology before finally presenting our results and conclusions.

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.

When modeling the decision of individuals to engage in crime, education has been relatively neglected by economists as a channel that might influence both criminal proceeds and the incentive to enter the illegal labor market. And yet, Ehrlich (1975) suggests that the relation between education and crime may 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.

Lack of individual data on criminal proceeds *and* education has prevented scholars from analyzing the effect of education on the productivity of criminals.[1] The only exception we are aware of is Carvalho and Soares (2013), 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 of years of education is not significantly different from zero. Instead, the coefficient on years of experience ranges between 5 and 10 percent.

[1] 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 of 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 US 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.[2]

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 of 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 OLS, 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, which 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

[2]On a related note, Lochner (2004) finds that education is associated with fewer property and violent crimes but with more white collar crimes (although not significantly).

the returns to schooling, focusing mainly on twin and sibling studies where estimates are based on within-family variation, and on instrumental variables (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 far 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 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 baseline 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 United States. The historical literature on education has traditionally focused on plotting the general trend of the rise of educational attainment and public education in the United States.[4] The general trend during the early twentieth century

[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 Claudia Goldin and Lawrence Katz, including Goldin and Katz (2008b) and Goldin and Katz (2008a).

was a steep upward trajectory in educational attainment associated with the "High School Revolution", with some states 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. The earliest estimates come from Goldin and Katz (2000), using the unique Iowa State Census of 1915, which provided information on wages and education levels on a large group of individuals. They found evidence of returns to a year of education as high as 9%, indicating that education was valuable even in agricultural states, and they further show that returns to education in Iowa fell from 1915 to 1940, with their estimates for the later date falling to around 6-7% for that state.

Other historical studies have focused on testing the impacts that compulsory schooling laws had on attainment and returns. Lleras-Muney (2002) used data from the 1960 Census to look at the effect of labor and schooling laws on the educational attainment of those who were aged 14 between 1915 and 1939. She finds that increasing the age at which a work permit could be obtained or lowering the school entrance age increased attainment on average by 5%. The main groups affected were those in the lower tail of the attainment distribution. Furthermore, states with higher concentrations of immigrants tended to pass more of these laws. These laws became better enforced from 1920 onwards, and have been analyzed again more recently, along with English-language schooling requirements in Lleras-Muney and Shertzer (2012), who focus on the effects on immigrants. They find a modest increase in literacy in English among the foreign-born and show that compulsory schooling laws had an effect on the foreign-born that was twice as big as that on natives, looking at the period 1910-1930 (these two papers imply that child labor laws working in conjunction with schooling laws were the key to explaining the closing gap between immigrant and native children's attainment, and can explain about three quarters of the increase in immigrant attainment overall, 1910-30). The average state's laws implied a minimum of 8 years of education for young working men in the 1940 Census.

Clay et al. (2012) look more comprehensively at compulsory schooling laws almost from their first use in 1880 up to 1927. 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 a year of education of 8-9%, using OLS, and 11-14% using IV methods where the compulsory schooling laws provide a plausible instrument. Our estimated returns for the control group of neighbors, reported below, are therefore 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 immigrant and Italian-American population 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.

Around the time of the unification of Italy in 1861 the Sicilian economy as well as the economy of the other southern Italian regions were highly dependent on four commodities: citrus, grain, wine, and sulfur. Sicily was the world leader in the production of sulfur and citrus, but around the turn of the twentieth century the price of these commodities plummeted (Buonanno et al., forthcoming).

The most serious blow was the discovery of sulfur in America which was considerably cheaper to extract. Moreover, the annexed south started to be heavily taxed under the new Italian central government. As a consequence of these, and other, adverse shocks, between 1901 and 1913 almost a quarter of Sicily's population departed for the United States. In that period around two million Italians, mainly from the South, emigrated to the US (Critchley, 2009). This massive wave of migration stopped with the passage of the U.S. Immigration Act of 1924.

The majority of these immigrants had been agricultural workers, with low levels of literacy. In Sicily children were expected to start working at an early age and schools were seen as a threat to the family's economic survival. In 1901 about 80 percent of the Sicilian population was illiterate (ISTAT, 2014), and such rates were likely to be 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 Federal Bureau of Narcotics data), was born in Chicago to a Sicilian-immigrant shoemaker and his Sicilian-immigrant wife. 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 in those days (Roemer, 1995). Later we will see 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 later became 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, like Lucky Luciano, Tommaso Lucchese, Vito Genovese, Frank Costello, etc., were children of these early immigrants. Criminal careers started quite early– Federal Bureau of Narcotics 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 (paradoxically) facilitated by Prohibition, a nationwide constitutional ban on the sale, production, importation, and transportation of alcoholic beverages that remained in place from 1920 to 1933, the second wave of immigrants that went on to become mafia bosses were already criminals by the time they entered the United States. Charles Gambino, Joe Profaci and others were in their 20s and 30s when they first entered the US, 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 and started adopting its code of honor and tradition.[5] In 1930 and 1931 these new arrivals sparked a mafia war, called the Castellammare war, named after a small city in Sicily where many of the new mafia bosses came from. The war lasted until Maranzano, who was trying to become the "boss of the bosses," was killed, probably by Lucky Luciano who had joined the Masseria Family.[6] This war put Lucky Luciano at the top of the mafia organization and there would be no more mafia wars during his government. Lucky Luciano died of heart attack at the airport of Naples in 1962.

By 1940, the Census year we use to collect information about education, income, housing value or rents of the mafia members listed in the 1960 FBN records, the mafia had a well established government, called "commissione." Joe Valachi, the first informant, revealed that the mafia, the *Cosa Nostra* ("our thing") was composed of approximately 25 Families. *Cosa Nostra* 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, a second in command, called underboss (*Sottocapo*), a counselor (*Consigliere*) and several captains (*Caporegime*) who headed a

[5] See Gosch and Hammer (1975).

[6] Before this event, in order to end the power-struggle between Masseria and Maranzano, Lucky Luciano had offered to eliminate Joe "the Boss" Masseria, which he did at an Italian restaurant by poisoning Masseria's food with lead.

group of soldiers (*regime*) (Maas, 1968).

These hierarchies allowed the mafia to successfully expand into a series of legal and illegal activities. Mobsters were involved in racketeering,[7] drug trafficking, gambling and bootlegging, but also owned restaurants, drugstores or were otherwise involved in the food 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:

> "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."

4 Data

In this section we explain how the dataset was constructed. We searched for 723 members of the Italian-American mafia whose details were listed in the 1960 Federal Bureau of Narcotics (FBN) records.[8]

[7]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.

[8]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

The records are an exact facsimile of the FBN secret files on American mafia members who were active in 1960 (MAF, 2007).[9]

We then link these records based on a multitude of variables (name, surname, names of family members, the residence address, the year of birth, arrests, etc.) *by hand* to the 1940 Decennial Census using the genealogical website 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.

Since between 1940 and 1960 mafia Families were not at war with each other, the second selection is likely to be more serious. Thirty-two mobsters out of 414 (7.7 percent) 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 of 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).

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.[10]

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 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

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).

[9]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.

[10]The results are available upon request.

U.S. Censuses, while Abramitzky et al. (2014) report a 19% match rate when connecting more than 1 census between 1900 and 1920.

We collected information about mafiosi and their closest neighbors, defined as all individuals recorded on the same page in the 1940 Census manuscript. 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 about highest grade of schooling attained, wage income, whether any non-labor income was earned in the previous year, migration in the past five years at the individual level and it also provides information on the house value or rent paid for each household. The resulting database on each mobster spans their criminal careers and early life history and background, 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 website, 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 organized crime families and their illegal behaviors to earlier information on educational attainment and family background has been possible. Mastrobuoni (forthcoming) provides more detail on the FBN source, but it contains information about 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 in the government (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 are also going to 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. In ninety percent of cases there are more than 25 such records. After the selection of mobsters and neighbors the average number of neighbors is equal to 6.2.

This gives us a final sample of 318 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 sample 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.

4.1 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 1 shows the summary statistics for mobsters and the matched neighbors.

Mobsters are considerably more likely to report no income, and the ones they report are on average 20 percent lower than for the 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. 31 percent), and their house is on average worth about 10 percent more compared to those of their neighbors. Moreover, even mobsters who are renting tend to spend 6 percent more than their closest matches. Since underreporting might bias our results we will conduct our analysis considering income, the value of their house and their monthly rent payments, as three different measures of their economic status.

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: 7.79 against 8.71.[11] In line with the anecdotal evidence mentioned before, Figure 2 shows that the difference is mainly driven by differences past secondary education, 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 21 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 smaller households. Geographic mobility is low for both groups: 85% of both mobsters and their neighbors have lived in the same house for at least five years.

Now that we have introduced the data sources and some summary statistics for our sample of mobsters and neighbors, the following sections will outline our methodology and results, facilitating a comparison of the returns to education in legitimate and illegimate activities.

[11] As a benchmark, the average years of education for U.S. males aged 25-59 from the 1940 Census as a whole was approximately 8.6 years (Goldin and Katz (2000), p. 790).

5 Counterfactuals, Returns to Education, and Returns to Experience

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.

These regressions are known to suffer from two main biases: measurement error bias and ability bias. The first refers to the fact that years of education might be measured with error. Since we do not use instruments it is important to address how these biases are likely to differentially influence estimates of the returns to education for mobsters and neighbors.

While we cannot assess the measurement error in education, we show that incomes are measured with more error among mobsters.

Mobsters are more likely to report incomes of zero (this is partly due to the fact that the 1940 Census did not require the self-employed to report their wage/salary income, and mobsters are more likely to be self-employed) and more likely to report low incomes compared to the value of their house or their rent. Given that a large part of mobsters' incomes are based on illegal activities this is not surprising. If similar misreporting happens with respect to education, returns to education would be biased towards zero.

If mobsters were more likely to misreport not just income but also education the, measurement error bias would be more severe among mobsters than among neighbors.

The ability bias is larger when the correlation between ability and education and the correlation between ability and income is stronger. Since mobsters are more likely to drop out of school for a series of reasons that are not related to ability (i.e. incarceration, but also cultural preferences. Later we will discuss historical evidence that Italian immigrant parents were forcing children to drop out of school at the age 14) the first correlation is likely to be weaker among mobsters than among their neighbors. As for the second correlation it is hard to say, but it would be difficult to make up a story where such abilities mattered more inside the mafia and at the same time more at school.

Summarizing the discussion, we believe the test of a negative difference between the returns to education in organized crime and those in the legal labor market to be a conservative one. And yet, we will see that such differences are estimated to be close to zero.

As is customary we transform all outcomes into logarithms (later we use alternative approaches to deal with zero incomes).

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 $\log \mu_i + \log Y_i = \alpha_i + y_i$ and can be captured by the constant term. 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.

The economic outcomes y_i^m (income, housing value, and rent) of mobsters ($m = 1$) and neighbors ($m = 0$) can be decomposed as a function of observables $w_i^m = (s_i^m, a_i^m, x_i^m)$, respectively education, age, and other observable characteristics, and unobservables $\theta^m = (\alpha^m, \delta^m, \gamma^m, \beta^m)$, respectively constant terms (including underreporting), returns to education, returns to experience (age), and returns from other observable characteristics.

Thus, in the most general terms we have that $y_i^m = f\left(w_i^m, \theta^m\right)$. Before adding more structure to $f\left(\cdot\right)$, we construct counterfactuals based on the propensity score $\Pr\left(m = 1\right) = \Phi\left(w_i'\zeta\right)$.

Following Abadie (2005), we weight observations by the (inverse) propensity score of assignment–i.e. the probability of belonging to each group, conditional upon the observed covariates. Specifically, the weight attached to each observation is:

$m_i + (1 - m_i)\,\frac{1-p}{p}\,\frac{\Phi\left(w_i'\zeta\right)}{1-\Phi\left(w_i'\zeta\right)}$ (multiplied by ω_i), where, p is the unconditional probability of being a mobster. When weighting we can construct the following counterfactual $\tilde{y}_i^0 = f(w_i^1, \theta^m)$, that is the outcome of neighbors if they had the same observable characteristics of mobsters. The difference $y_i^1 - \tilde{y}_i^0 = f(w_i^1, \theta^1) - f(w_i^1, \theta^0)$ measures the importance of unobservable characteristics, while $\tilde{y}_i^0 - y_i^0 = f(w_i^1, \theta^0) - f(w_i^0, \theta^0)$ measures the importance of observable characteristics.

Adding these two terms together one obtains the original differences between the outcomes of mobsters and neighbors.[12]

We estimate the propensity score using a probit regression for the probability of being a mobster, conditional on all the observable characteristics listed in Table 1. The results are shown in Table 2. Education is associated with a lower probability of becoming a mobster and so is age. Being born in Italy is a strong predictor of becoming a mobster.

Figure 1 shows the distribution of the propensity score for the two groups. While the density for the mobsters is significantly shifted toward a probability of one, the two densities do overlap, meaning that for almost all propensities scores one can find mobsters as well as neighbors.

Using the propensity score we build the counterfactual distributions of neighbors if they had the same observable characteristics of mobsters. These counterfactuals are presented in Figures 3 to 5, together with the cumulative distributions of log income, log housing value, and log rent of mobsters and neighbors.

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 values or rents. This is likely driven by income being underreported.

If observable characteristics were able to explain these differences between mobsters and neighbors, the counterfactual distribution of neighbors would lie on top of the one of mobsters. If anything, observable characteristics accentuate the differences. Neighbors

[12]One important advantage of propensity score weighting, relative to other matching estimators, is the possibility of computing asymptotically valid standard errors by bootstrapping methods (Abadie and Imbens, 2008).

with the observable characteristics of mobsters would earn the same income they already earn when incomes are below the median. But for incomes above the median they would actually earn even less than mobsters.

This is even more true when looking at housing values and rents. Despite the lower reported incomes mobsters live in more expensive housing and pay higher rents. If neighbors had the same characteristics of mobsters, for example lower education, they would live in even less expensive housing arrangements.

Next we try to measure the returns to education for mobsters and neighbors. In order to estimate such returns it is customary to add more structure to $f(\cdot)$. We follow the long tradition of Mincer regressions and use linear models, where the log of y is regressed on years of education and age. [13]

The Mincer regressions for the two groups are:

$$\log y_i^m = \alpha^m + \theta^m w_i^m + \epsilon_i^m.$$

Similarly to the counterfactual regressions, differences in y can be decomposed as follows:

$$\Delta \log y_i = \Delta\alpha + \Delta\theta w_i^1 + \theta^0 \Delta w_i + \Delta\epsilon_i, \tag{1}$$

where $\Delta z_i = z_i^1 - z_i^0$.

Due to lack of direct evidence on workplace experience, in all Mincer wage regressions we control for age as opposed to experience. While experience is often proxied by age minus the years of education, in our data a non-negligible fraction of individuals have very few years of education (12 percent have less than 4 years of education). In order

[13] 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.

to compute the return to education keeping experience rather than age constant it is sufficient to add the returns to age to the returns to education.

The results for the Mincer regressions of income when controlling only for age and years of education are in Column 1 and 5 of Table 4 for mobsters and neighbors. The returns to education and to experience are equal to about 7 and 4 percent, and do not differ for the two groups. This result does not change when controlling for a set of socioeconomic factors (columns 2 and 6), for state fixed effects (columns 3 and 7), and for city fixed effects (columns 4 and 8). The robustness table 5 shows that the results differ little when controlling for a squared term of age (Column 2), for log of total working hours per year (a variable that is not always available), and when limiting the sample to neighbors whose age gap with mobsters is under 5 years (and not 10 as in the baseline). The results are also robust to focussing on individuals above age 22, an age at which individuals in the 1940s had usually completed their studies. The returns to education for mobsters stop being significant when using a Poisson model, i.e. when including incomes of zero. 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.

Since income is subject to these biases, we also estimate Mincer regressions using housing outcomes of head of households as dependent variables. Housing outcomes might be an alternative, possibly less distorted, proxy for long-term income or wealth. Table 6 shows that even when using housing outcomes the returns to education and experience are quite comparable for the two groups. If anything, the returns to education seem to be larger for mobsters than for neighbors.

That the differences between the returns to education are not significantly different from each other can be seen when performing the Oaxaca decomposition of Equation 1. Table 7 decomposes the differences between mobsters and neighbors in explained (Δw)

and unexplained ($\Delta\theta$) differences.

6 Conclusions

This paper analyzes the link between education and economic outcomes (income, housing values and housing rents) among members of the Italian-American Mafia and among their neighbors. We focus on mobsters who were listed in a 1960 Federal Bureau of Narcotics 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 statistically 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. While this might appear counterintuitive, a model of human capital investment where the working life (in this case because of expected prison time, injuries or death) is shorter, predicts larger returns to education.

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

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, etc. might indeed increase with education and be necessary for success in these mafia roles. Moreover, often times loan sharking would allow mobsters to acquire legitimate businesses which would also be convenient for money-laundering purposes. Returns to education in such activities are likely to be large too.

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.

References

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

Alberto Abadie. Semiparametric difference-in-differences estimators. *Review of Economic Studies*, 72(1):1–19, 2005.

Alberto Abadie and Guido W Imbens. On the failure of the bootstrap for matching estimators. *Econometrica*, 76(6):1537–1557, 2008.

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

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

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

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

P. Buonanno, R. Durante, G. Prarolo, and P. Vanin. Poor institutions, rich mines: Resource curse and the origins of the sicilian mafia. *The Economic Journal*, forthcoming. forthcoming.

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

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

Leandro Carvalho and Rodrigo Soares. Living on the edge: Youth entry, career and exit in drug-selling gangs. Technical report, IZA Discussion Paper, 2013.

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

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

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

Isaac Ehrlich. On the relation between education and crime. In *Education, Income, and Human Behavior*, pages 313–338. NBER, 1975.

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

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

Claudia Goldin and Lawrence Katz. *The race between education and technology*. Belknap Press of Harvard University Press, Cambridge, MA, 2008a.

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

Claudia Goldin and Lawrence F Katz. Mass secondary schooling and the state: The role of state compulsion in the high school movement. *NBER Chapters*, pages 275–310, 2008b.

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

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

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

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

Adriana Lleras-Muney. Were compulsory attendance and child labor laws effective? an analysis from 1915 to 1939*. *Journal of Law and Economics*, 45(2):401–435, 2002.

Adriana Lleras-Muney and Allison Shertzer. Did the americanization movement succeed? an evaluation of the effect of english-only and compulsory schools laws on immigrants' education. Technical report, National Bureau of Economic Research, 2012.

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

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

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

Giovanni Mastrobuoni. The value of connections: Evidence from the italian-american mafia. *Economic Journal*, forthcoming.

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

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

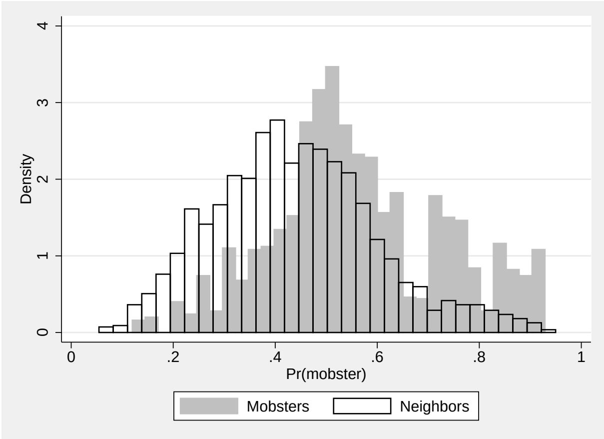

Figure 1: Histogram of the Propensity Score

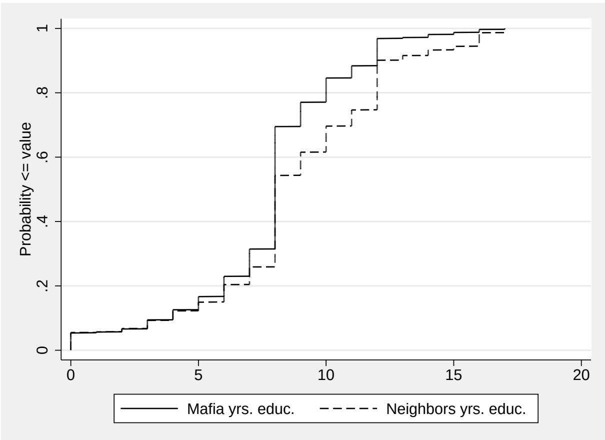

Figure 2: Cumulative Distribution of Years of Education

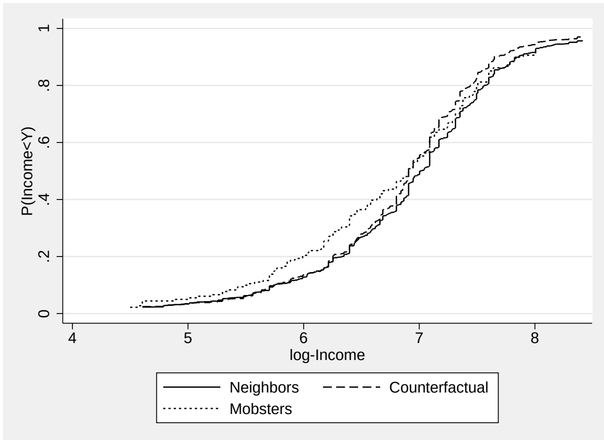

Figure 3: Cumulative Distribution of Years of (log) Income

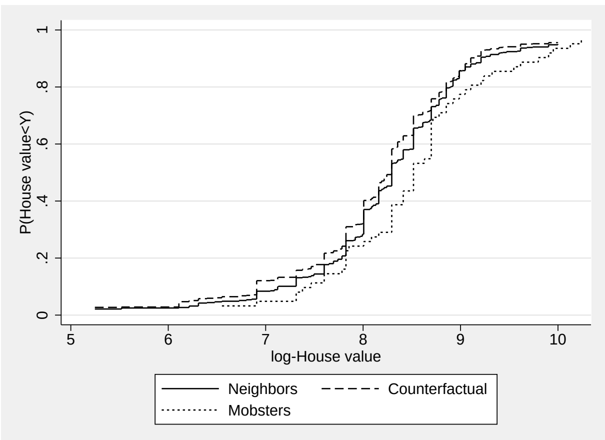

Figure 4: Cumulative Distribution of (log) Housing Value

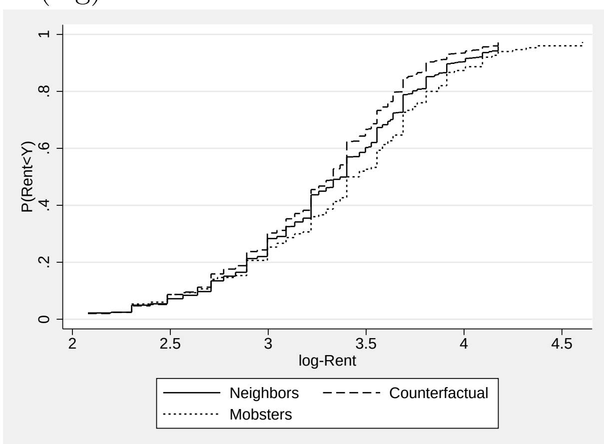

Figure 5: Cumulative Distribution of (log) Rent

Table 1: Summary statistics

	Mafia Mobsters			Neighbors		
	Mean	St. Dev.	N	Mean	St. Dev.	N
Income	817	1,164	292	1,067	1,178	1783
Zero income	0.38	0.49	292	0.23	0.42	1783
Head of HH	0.67	0.47	318	0.58	0.49	1973
Home owner	0.33	0.47	318	0.31	0.46	1973
House value	7,336	10,188	104	6,648	10,314	533
Rent	35.74	25.29	201	33.71	27.67	1272
Yrs. of education	7.79	3.08	318	8.71	3.71	1973
Age in years	33.77	8.44	318	34.11	9.76	1973
Married	0.72	0.45	318	0.63	0.48	1973
Born in Italy	0.38	0.49	318	0.15	0.36	1973
Alien citizen	0.30	0.46	318	0.21	0.41	1973
Naturalized citizen	0.25	0.43	318	0.18	0.39	1973
# of HH members	2.84	2.15	318	3.38	2.03	1973
# of children	0.50	1.02	318	0.75	1.21	1973
Same residence last 5yrs.	0.84	0.37	318	0.85	0.36	1973

Table 2: Propensity Score Probit

	(1)	(2)
	P(mobster=1)	
	Coeff.	Marginal
Yrs. of education	-0.033***	-0.013***
	(0.011)	(0.004)
Age in years	-0.029***	-0.012***
	(0.004)	(0.002)
Married	0.271***	0.108***
	(0.090)	(0.035)
Born in Italy	1.243***	0.448***
	(0.123)	(0.037)
Alien citizen	-0.564***	-0.220***
	(0.188)	(0.070)
Naturalized citizen	0.142	0.057
	(0.175)	(0.070)
# of HH members	-0.045	-0.018
	(0.029)	(0.012)
# of children	-0.109**	-0.044**
	(0.055)	(0.022)
Same residence last 5 yrs.	-0.005	-0.002
	(0.090)	(0.036)
Constant	1.067***	
	(0.191)	
Observations	3946	3946
log-likelihood	-396.5	

Notes: propensity score with clustered (by mobster) standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Correlations Between Measures of Wealth and Income

	(log) House value	(log) Rent
Whole Sample		
Mobster Income	0.267	0.234
Mobster log-Income	0.514	0.340
Neigh. Income	0.291	0.397
Neigh. log-Income	0.334	0.435
Heads of Household		
Mobster Income	0.280	0.236
Mobster log-Income	0.641	0.441
Neigh. Income	0.512	0.433
Neigh. log-Income	0.452	0.552

Notes: Correlation coefficients

Table 4: 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	0.072***	0.068**	0.061**	0.063**	0.081***	0.077***	0.076***	0.074***
	(0.026)	(0.028)	(0.027)	(0.029)	(0.008)	(0.008)	(0.008)	(0.009)
Age in years	0.044***	0.043***	0.037**	0.033**	0.039***	0.029***	0.029***	0.029***
	(0.014)	(0.016)	(0.015)	(0.016)	(0.004)	(0.004)	(0.004)	(0.005)
Married		0.219	0.174	0.232		0.483***	0.469***	0.463***
		(0.164)	(0.160)	(0.188)		(0.075)	(0.071)	(0.066)
Born in Italy		-0.478	-0.614	-0.878		-0.040	-0.056	-0.043
		(0.378)	(0.393)	(0.564)		(0.107)	(0.103)	(0.103)
Alien citizen		0.912**	1.146**	1.339*		-0.062	-0.085	-0.041
		(0.457)	(0.469)	(0.704)		(0.102)	(0.102)	(0.108)
Naturalized citizen		-0.672**	-0.775***	-0.773**		0.037	0.055	0.016
		(0.281)	(0.282)	(0.346)		(0.106)	(0.107)	(0.116)
# of HH members		-0.090	-0.141**	-0.186**		-0.008	-0.021	-0.027
		(0.067)	(0.070)	(0.083)		(0.021)	(0.020)	(0.020)
# of children		0.161	0.232*	0.352**		-0.028	-0.010	0.003
		(0.133)	(0.132)	(0.161)		(0.032)	(0.031)	(0.032)
Same residence last 5 yrs.		-0.049	0.030	0.080		0.072	0.057	0.013
		(0.214)	(0.200)	(0.289)		(0.089)	(0.082)	(0.085)
Constant	4.716***	4.909***	5.225***	5.369***	4.844***	4.898***	4.976***	5.027***
	(0.578)	(0.762)	(0.752)	(0.857)	(0.161)	(0.205)	(0.193)	(0.208)
State fixed effects			Y				Y	
City fixed effects				Y				Y
Observations	1143	1143	1143	1143	1381	1381	1381	1381
R-squared	0.096	0.148	0.268	0.471	0.206	0.265	0.298	0.383

Notes: There are a total of 318 mobsters in the data. The number of observations refer to the unweighted data. Weighting there are 318 observations in both groups. Mincer wage regressions with clustered (by mobster) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Robustness Mincer Wage Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
	log-Income					
Panel A: Mafia Members	*Baseline*	*w. Age sq.*	*w. Tot.Hrs.*	$\Delta age \leq 5$	$age \geq 22$	*Poisson*
Mafia yrs. educ.	0.068**	0.065**	0.061***	0.092***	0.056**	0.045
	(0.028)	(0.028)	(0.022)	(0.024)	(0.028)	(0.029)
Age in years	0.043***	0.047***	0.024*	0.058***	0.039**	0.032**
	(0.016)	(0.014)	(0.013)	(0.015)	(0.017)	(0.015)
Age squared		-0.002				
		(0.002)				
log-Total hours worked			0.798***			
			(0.134)			
Observations	1,143	1,143	779	684	1,071	1,788
R-squared	0.148	0.161	0.507	0.227	0.113	
log-likelihood						-183081
Panel B: Neighbors	*Baseline*	*w. Age sq.*	*w. Tot.Hrs.*	$\Delta age \leq 5$	$age \geq 22$	*Poisson*
Neighbors yrs. educ.	0.077***	0.074***	0.062***	0.075***	0.076***	0.072***
	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)	(0.010)
Age in years	0.029***	0.037***	0.022***	0.038***	0.022***	0.023***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)
Age squared		-0.002***				
		(0.000)				
log-Total hours worked			0.558***			
			(0.057)			
Observations	1,381	1,381	1,124	814	1,279	1,783
R-squared	0.265	0.294	0.411	0.272	0.202	
log-likelihood						-140536

Notes: There are a total of 318 mobsters in the data. The number of observations refer to the unweighted data. Weighting there are 318 observations in both groups. Mincer wage regressions with clustered (by mobster) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Mincer Regressions Using House Values and Monthly Rents

	(1)	(2)	(3)	(4)	(5)	(6)
	Mafia members			Neighbors		
	Home-owner	House-value	Rent	Home-owner	House-value	Rent
Yrs. of education	0.020	0.086*	0.050***	0.034**	0.059**	0.051***
	(0.030)	(0.045)	(0.018)	(0.017)	(0.023)	(0.009)
Age in years	0.088***	0.044**	0.027**	0.054***	0.002	0.020***
	(0.017)	(0.021)	(0.010)	(0.008)	(0.011)	(0.005)
Other Xs	Y	Y	Y	Y	Y	Y
Observations	1,228	341	876	1,080	252	811
R-squared	.	0.154	0.136	.	0.155	0.167
log-likelihood	-109.6			-95.77		

Notes: There are a total of 318 mobsters in the data. The number of observations refer to the unweighted data. Weighting there are 318 observations in both groups. Mincer wage regressions with clustered (by mobster) standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Blinder-Oaxaca Decomposition

Differential in	Coef. log-Income	Std. Err. log-Income	Coef. log-Housevalue	Std. Err. log-Housevalue	Coef. log-Rent	Std. Err. log-Rent
Neighbors' Xb	6.91	0.04	8.31	0.10	3.37	0.04
Mobsters' Xb	6.75	0.08	8.51	0.14	3.42	0.05
Difference	0.16	0.08	-0.20	0.14	-0.04	0.04
Explained difference between Neighbors and Mobsters: Δw						
Yrs. of education	0.09	0.02	0.12	0.05	0.05	0.02
Age in years	0.04	0.01	-0.03	0.02	0.01	0.01
Married	-0.03	0.02	-0.03	0.05	-0.01	0.01
Born in Italy	0.01	0.02	0.12	0.12	0.01	0.02
Alien citizen	0.00	0.01	-0.09	0.11	0.00	0.00
Naturalized citizen	0.00	0.01	0.05	0.09	0.00	0.01
# of HH members	0.00	0.01	0.02	0.05	0.03	0.02
# of children	-0.01	0.01	-0.02	0.03	-0.03	0.01
Same residence last 5 yrs.	0.00	0.00	0.00	0.00	0.00	0.00
Total	0.09	0.03	0.15	0.12	0.05	0.03
Unexplained difference between Neighbors and Mobsters: $\Delta\theta$						
Yrs. of education	0.06	0.22	-0.12	0.30	-0.01	0.14
Age in years	-0.46	0.49	-1.12	0.86	-0.29	0.32
Married	0.19	0.12	-0.32	0.44	-0.53	0.16
Born in Italy	0.14	0.12	-0.01	0.27	0.02	0.07
Alien citizen	-0.26	0.13	-0.05	0.32	-0.04	0.06
Naturalized citizen	0.16	0.07	0.10	0.25	0.05	0.05
# of HH members	0.22	0.18	0.68	0.32	0.15	0.12
# of children	-0.08	0.06	-0.31	0.18	-0.08	0.04
Same residence last 5 yrs.	0.10	0.18	-0.15	0.52	-0.04	0.10
Constant	-0.01	0.74	0.96	1.39	0.66	0.40
Total	0.07	0.08	-0.35	0.19	-0.10	0.04
Number of obs.	2,524		689		1,739	

Notes: Clustered (by mobster) standard errors