Draft Extended Abstract: Please Do Not Quote, Cite, or Circulate

The Departed: Origins and Persistence of Organized Crime in the United States

Massimo Anelli, Paolo Pinotti, Zachary Porreca
Bocconi University

November 2024

Abstract

1 Introduction

Organized crime groups are among the most enduring organizations in human history. The origins of groups such as the Mafia, Camorra and 'Ndrangheta, some of the oldest and most notorious criminal organizations worldwide, can be traced back to the formation of the Italian state in the mid-19th century (Gambetta, 1992). This historical link has led many to argue that organized crime is deeply rooted in local culture (e.g., Schneider and Schneider, 1994), which exhibits significant persistence over time (Guiso et al., 2016). A key implication of this "cultural hypothesis" is that organized crime would struggle to successfully transplant itself into different cultural contexts.

From the mid-20th century onward, however, Italian criminal organizations expanded their influence beyond their traditional strongholds in Sicily, Campania, and Calabria, extending their reach into other Southern Italian regions and eventually into the wealthier Northern regions.On the other hand, evidence on the transplantation of criminal organizations to entirely different countries remains limited. Studies by Sviatschi (2022) and Melnikov et al. (2024) examine the deportation of thousands of U.S.-based immigrants with criminal backgrounds—including members of the MS-13 gang—back to Central America, where their presence led to increases in homicides,

gang activity, and reduced employment opportunities. However, these gang members returned to countries sharing their cultural and linguistic backgrounds, making their cases somewhat unique.

In this paper, we focus on the forced migration of the Sicilian Mafia in the 1920s from one of Europe's poorest and most insular regions to the United States—a country vastly different from Sicily in terms of social norms, language, and formal institutions, and at the time, one of the richest nations in the world. Specifically, we trace the origins of the Italian-American Mafia, commonly known as La Cosa Nostra (LCN), to the arrival of hundreds of prominent Sicilian Mafia members fleeing repression by the Fascist regime.

In an effort to eliminate a powerful rival to his regime in Sicily, Mussolini entrusted Prefect Cesare Mori with leading an extensive military campaign against the Mafia. On January 1, 1926, Mori laid siege to the town of Gangi, a known hub for criminal groups and bandits. Over the following months, he raided another 31 Mafia strongholds, arresting or killing numerous members and forcing many others to flee Sicily and join previous waves of Italian migrants in the United States. Despite the restrictive quotas imposed by the 1924 Immigration Act, the President's Commission on Organized Crime (1984) estimates that at least 500 members of the Sicilian Mafia found safe haven—and indeed, new criminal opportunities—in the United States. Although the sources and methods behind these estimates are unclear, notable Mafia clans, such as the Gambinos, Buonannos, and Corleones, are cited in other sources (see, e.g., Lupo, 2008).

Migration literature consistently shows that new arrivals tend to settle in enclaves of previous migrants from their home country, often from the same villages or towns (see, e.g., Card 2001, and Munshi, 2003). Using detailed information from full-count Census data, maps of enumeration districts, and Social Security mortality files, we divide eleven of the largest U.S. cities into over 8000 time-invariant hexagons of approximately 1.67 square kilometers each. Within these hexagons, we identify nearly 500 that in 1920 hosted Sicilian migrants from the 32 towns raided by Mori (henceforth referred to as "Mori neighborhoods"). Using additional digitized FBI data on Italian-American Mafia activity in 1959, we then compare the spatial distribution of Mori neighborhoods with the 1920 geocoded residences of men who would later become Mafia bosses in 1959, as well as the geocoded locations of their criminal activities.

Our findings reveal that 75% of Mafia bosses in 1959 had resided in Mori neighborhoods (representing only 1.7% of all hexagons) back in 1920. This overlap rises to 88% when the analysis is limited to New York, Chicago, and Philadelphia—the cities with the largest Mafia presence. Drawing 1,000 random samples of placebo hexagons,

we find that the probability of achieving the same matching rates is effectively zero. Notably, this p-value remains zero even when sampling placebo hexagons exclusively within those hosting Sicilian immigrants from towns not raided by Mori. Although the share of Mafia activities located in Mori neighborhoods is somewhat lower, it remains significant—accounting for approximately one-third of all Mafia activities in 1959, with a p-value of 0.

These results confirm that Mafia associates fleeing Sicily provided a critical source of "criminal capital," which proved fundamental to the rise of the Italian-American Mafia in the following decades. We then examine the socio-economic consequences of Mafia presence, uncovering contrasting patterns in the short and long term.

Beginning with short-term effects, we find that by the late 1930s, Mori neighborhoods were significantly more likely than other neighborhoods to be "redlined" by the Home Owner's Loan Corporation, the federal agency responsible for assessing credit and investment risk in urban neighborhoods. While it is not particularly surprising that areas with higher immigrant populations were subject to redlining, Mori neighborhoods were 40% more likely to be redlined even when compared to other neighborhoods with Sicilian migrants. This differential can be attributed to the greater presence of criminal groups in Mori neighborhoods, which was explicitly noted in some assessments as a reason for lower ratings.

A substantial body of literature documents the negative long-term effects of redlining on various socioeconomic dimensions, including lower rates of home ownership, lower housing values, lower rent levels, and higher rates of vacancy across subsequent decades (Aaronson et al., 2021). Remarkably, we find that by the 2000s-2010s, Mori neighborhoods exhibited *better* outcomes than other areas, with higher shares of college graduates and higher employment rates. Although unexpected given the well-documented adverse consequences of redlining, these positive long-term effects are consistent with previous findings by Murphy and Rossi (2020) regarding the impact of drug cartel presence across Mexican municipalities.

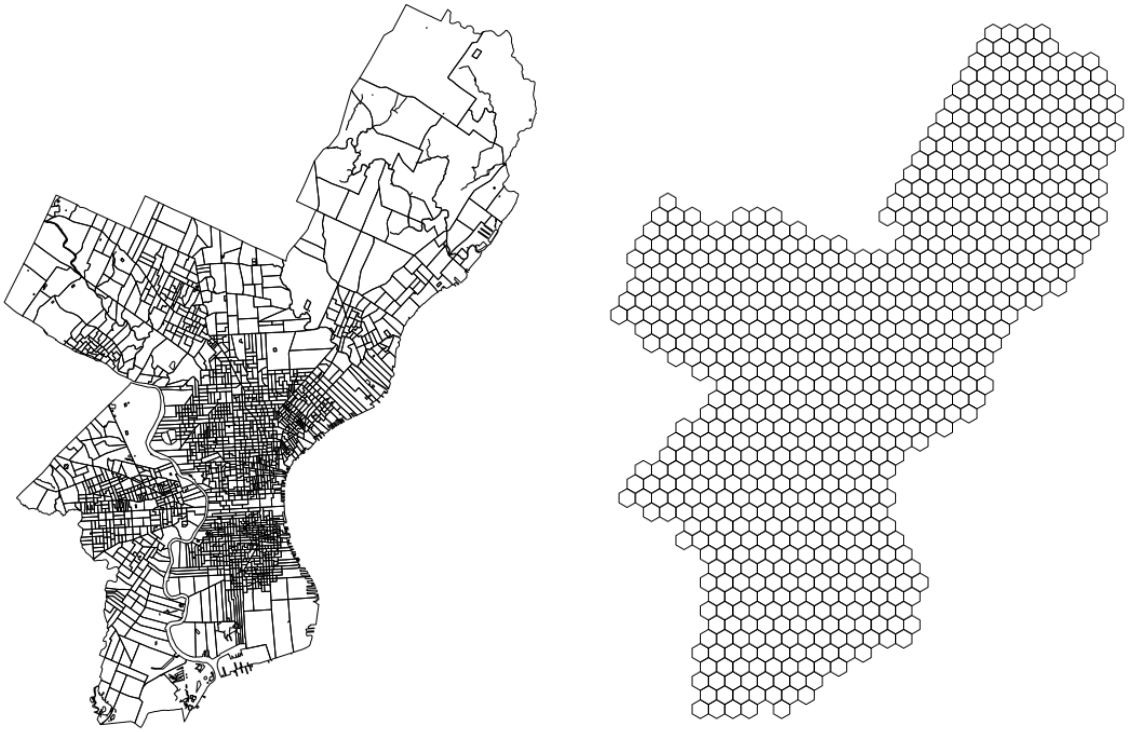
2 Data

To address these questions related to the migration of the Sicilian mafia to America, we leverage a wealth of historical and contemporary data. Primarily, we employ both restricted and unrestricted versions of the full count 1900-1940 US censuses, made available through IPUMS, the Social Security Administration's Numerical Identification System Mortality Files (Numident), made available by NARA, maps of historical Census enumeration districts provided by the Urban Transition Historical

GIS Project (Logan et al., 2015), historical redlining maps made available by the Mapping Inequality project (Nelson et al., 2023), digitized records of an FBI report on the US mafia (Mastrobuoni 2015), and additional data provided by Opportunity Insights (Chetty et al., 2018).

Our primary strategy entails employing a novel cross-walking algorithm to identify the Census records of individuals in the Numident records born in identifiable Sicilian commune targeted by Mori's raids. This algorithm allows for name misspellings and anglicizations, slight discrepancies in age, and, when available, leverages information on the names of parents. Further details of the matching algorithm can be found in the appendix. Overall, we achieve a 17% match rate, comparable to the match rate from most same-individual across-census algorithms. These individuals, along with other Sicilian-born individuals, are mapped to their home enumeration districts and are tracked across time using the new Census Tree data and algorithm, which itself achieves up to an 89% match rate (Buckles et al. 2024). These enumeration districts are then layered into temporally-constant equally sized hexagons, as in Shertzer et al. (2016). These hexagons are roughly 1.67 square kilometers in area and are meant to be comparable to modern Census tracts. They serve as the units of analysis in which all of the paper's results will focus. In total we are able to replicate this strategy for eleven of the US's largest cities of the era: Baltimore, Boston, Brooklyn, Chicago, Cincinnati, Cleveland, Detroit, Manhattan, Philadelphia, Pittsburgh, and St. Louis.

Figure 1, as an example, depicts the 1920 enumeration district map of Philadelphia on the left and the corresponding grid of equally sized hexagons that we map this to to employ as our primary units of analysis.



((a)) Enumeration district map ((b)) Hexagon grid map

Figure 1: The map on the left depicts the 1920 Census enumeration districts of Philadelphia. The map on the right depicts the grid of equally sized hexagons that we overlay to use as our primary units of analysis.

Our first set of results, documenting the emergence of the American arm of the Sicilian mafia, La Cosa Nostra, in the enclaves of immigrants from the small set of the 32 villages targeted by Cesare Mori relies on the digitized records of a 1959 FBI report on the American mafia, previously employed in Mastrobuoni (2015). We first replicate our Numident matching algorithm for the set of 810 mafiosi identified in the report- identifying the 1910 and 1920 residence locations of 17% or 142 of the recorded 1959 mafiosi. We then geocode all addresses associated with the mafia of the era, assigning those addresses to their constituent hexagons. We proceed by comparing the spatial distribution of *future* mafiosi/mafia locations to that of 1920 (pre-Mori raids) enclaves of immigrants from commune Mori would target in the subsequent decade. This procedure is detailed in a subsequent section.

Subsequent analysis data sets are constructed by mapping other spatial data onto our hexagon maps. First, we overlay the Home Owners' Loan Corporation maps of

the late 1930s (so called "redlining maps") with our hexagon grids, to test for differential rates of exclusion from the formal financial sector for "Mori" neighborhoods. Effectively, we seek to test whether in the short term the banking sector capitalized mafia-criminality into their neighborhood assessments. We proceed by overlaying contemporary zip code and Census tract maps over the hexagons to merge in data from Opportunity Insights to allow for the analysis of long run contemporary outcomes.

In the next section we outline the details of each empirical strategy employed, alongside results.

3 Empirical Strategies and Results

3.1 Spatial Distribution of Future Mafia Activity

To begin, we provide evidence that the American mafia of 1959 was inextricably linked spatially to the locations of early immigrants from the Sicilian commune targeted by Mori's raids, before those raids occurred. Mori's raids were conducted in the latter half of the 1920s. As such, in this exercise we focus on the 1920 distribution of immigrants in the US originating from the commune his forces would later target. Card (2009) laid the foundation for what is now a seminal result in the economics of migration: later waves of immigrants tend to settle among prior immigrants from their home area. Historians and law enforcement have suggested that mafia members fled Sicily in response to the raids and entered the United States to form the roots of the American LCN. Thus, if this is true, we should observe a strong relationship between the pre-raid locations of "Mori commune" immigrants and locations later associated with the mafia.

Here, we have two spatial distributions that we are interested in relating to these 1920 immigrant enclaves. First, we are interested in the 1920 locations of the men who would grow to become *known* mafia bosses by 1959. Second, we are interested in the spatial distribution of locations associated with mafia activity in 1959.

The first of these distributions provides us with a test of the migration of criminal capital hypothesis. The locations of mafia bosses 40 years prior largely corresponds to their childhood or young adult residences. If we are to observe future mafiosi disproportionately arising from the same neighborhoods where the old world mafia was likely to settle when displaced from Sicily, we can attribute this to the shock of an arrival of criminal capital in the 1920s providing the infrastructure and career paths which were able to grow these men into mafia bosses decades later.

The second distribution of interest, the known locations of mafia activity from the 1959 mafia report, is able to provide insights into the locational preferences of the mafia. Organized crime bestows benefits upon the communities in which it locates (Murphy & Rossi 2020). Arguably these benefits are bestowed as a means of incentivizing local cooperation with their activities (Benzecry et al. 2024). It is intuitive to assume that due to this "symbiotic" relationship between organized crime and local communities, that the operations of mafias and similar criminal actors would concentrate within the communities in which they arise form, where support is initially strongest.

To test the relationships between these spatial distributions and the 1920 locations of "Mori commune" immigrants in America, we employ a test of randomization inference. We map immigrants, 1959 mafiosis' matched 1920 Census records, and geocoded locations of 1959 mafia activity to their constituent hexagons. After calculating both the proportion of future mafia bosses' 1920 residences and 1959 known mafia locations that are in the hexagons containing immigrants originating in "Mori commune" in Sicily prior to Mori's raids, we construct an empirical p-value for each proportion through randomization inference. We draw random placebo sets of "Mori immigrant hexagons" and calculate the proportion of future mafiosi locations or future mafia activity locations present within that set. After iterating through 1000 randomized placebo samples, we are able to calculate an empirical p-value from distribution of placebo proportions. We then proceed to replicate this exercise on progressively stricter super-sets from which our placebo samples are drawn.

We begin with the universe of all hexagons in our 11 cities, then proceed to limit our sample to only above-median Italian-born proportion districts, districts with identifiable (from our Numident crosswalk) Sicilian born immigrants, districts that are above-median Italian with identifiable Sicilians, and lastly we replicate all of those analysis limited to the cities with the most identifiable 1959 mafia activity (Brooklyn, Chicago, Manhattan, and Philadelphia). Table 1 depicts all results from this randomization exercise. Figures 2 and 3 visually depict the results from these exercises corresponding to limiting our random sampling to the neighborhoods with the densest proportions of Italian-born immigrants that also contain Sicilian-born immigrants in the three cities that housed the most mafia activity in 1959.

		Locations of 1959 Mafia Activity		Residences of Future 1959 Mafiosi in 1920	
		Observed Prop.	p value	Observed Prop.	p value
All disticts	All cities	0.357	0.00	0.750	0.00
	New York City, Chicago, and Philadelphia	0.331	0.00	0.880	0.00
	New York City and Chicago	0.317	0.00	0.872	0.00
Above median Italian immigrant districts	All cities	0.334	0.00	0.736	0.00
	New York City, Chicago, and Philadelphia	0.366	0.00	0.780	0.00
	New York City and Chicago	0.325	0.00	0.745	0.00
Districts with Sicilian immigrants	All cities	0.280	>0.01	0.750	0.00
	New York City, Chicago, and Philadelphia	0.418	0.00	0.880	0.00
	New York City and Chicago	0.399	0.00	0.870	0.00
Above median Italian immigrant districts with Sicilian immigrant presence	All cities	0.326	0.00	0.736	0.00
	New York City, Chicago, and Philadelphia	0.354	0.02	0.950	0.00
	New York City and Chicago	0.313	0.43	0.744	0.00

Table 1: Table depicting all results from randomization exercise.

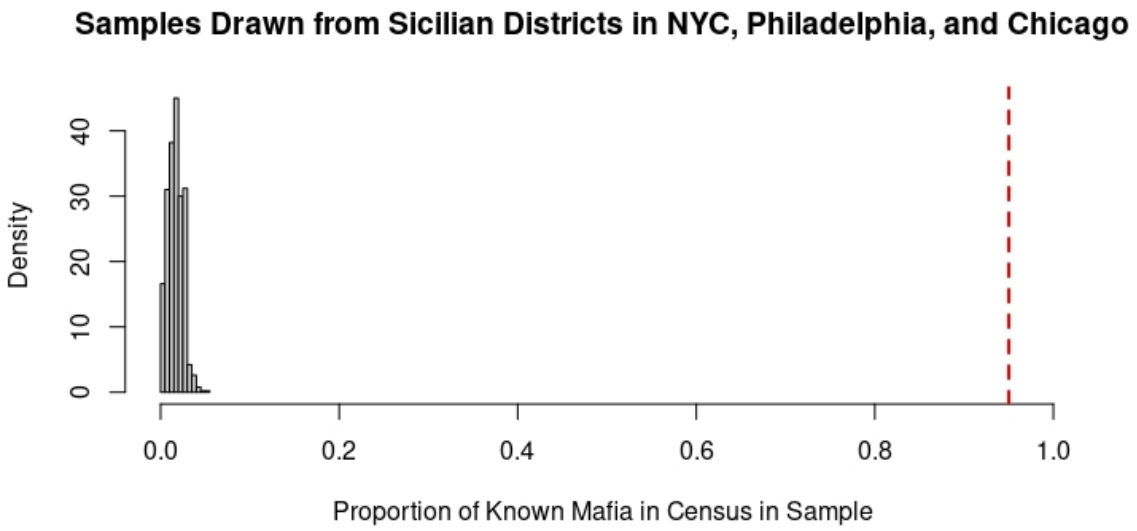


Figure 2: Figure depicting the observed proportion (the red dashed line) of known 1959 mafia members, that were able to be matched to the 1920 Census, that lived in neighborhoods housing immigrants from "Mori commune" in 1920. The distribution to the left depicts the proportion of those mafiosi living in randomly drawn samples of neighborhoods drawn from the most-Italian neighborhoods that also housed Sicilian immigrants in 1920. For this figure, the sample is limited to the three cities that were home to the most mafia activity and mafia members in 1959.

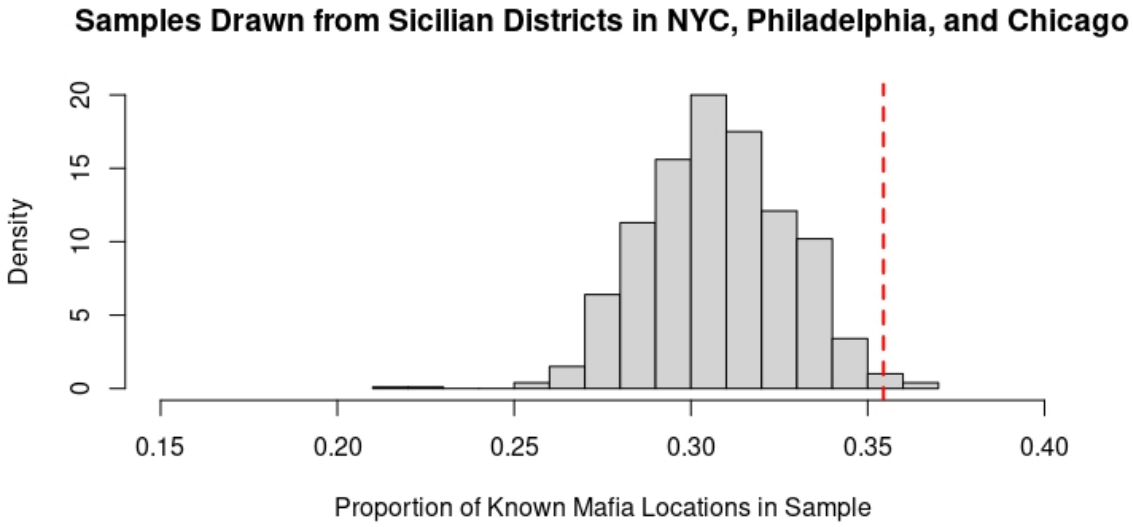


Figure 3: Figure depicting the observed proportion (the red dashed line) of geocoded locations of mafia activity in 1959 that were in neighborhoods housing immigrants from "Mori commune" in 1920. The distribution to the left depicts the proportion of those geocoded locations that are in randomly drawn samples of neighborhoods drawn from the most-Italian neighborhoods that also housed Sicilian immigrants in 1920. For this figure, the sample is limited to the three cities that were home to the most mafia activity and mafia members in 1959.

Figures 2 and 3 and Table 1 provide evidence of a strong linkage between the pre-raid locations of "Mori immigrants" in US cities and the spatial distribution of the American mafia 40 years later. We interpret this as evidence of an inextricable link between the mafia in America and the 32 Sicilian commune from which those immigrants originated. The men who would come to be in power of the mafia years later disproportionately arose from the enclaves of those immigrants- with observed proportion far exceeding anything close to what randomization could provide, even with the strictest sampling restrictions. This result is reminiscent of Sviatschi (2022), where it was demonstrated that youth in El Salvador who were exposed to the arrival of deported US gang members were more likely to engage in gang activity themselved later on. Similarly, the locations associated by the FBI with the American mafia of 1959 were also disproportionately located within those same neighborhoods that housed pre-raid "Mori" immigrants.

3.2 Redlining

Here we seek to address whether during the period following Mori's raids in Sicily, and the arrival of expelled mafiosi in America, contemporary financial institutions capitalized concerns regarding criminality into their assessments of the neighborhood enclaves of Mori commune immigrants. In the late 1930's a federal agency called the Home Owner's Loan Corporation (HOLC) classified urban neighborhoods across the US based off of their relative riskiness of investment (Aaronson et al., 2021). These classifications and the maps based on them famously incorporated race and immigration characteristics into their underlying assessments (Fishback et al., 2020). HOLC maps were a strong influence on (the largely no longer surviving) Federal Housing Administration maps, which were directly used to assess the credit-worthiness of insured loans (Light 2010). Neighborhoods marked as degraded were explicitly barred from having mortgages underwritten (Light 2010). The maps, and their impacts on lending behavior from formal financial institutions led to long lasting adverse housing market outcomes, namely lower rates of home ownership, lower housing values, lower rent levels, and higher rates of vacancy across subsequent decades (Aaronson et al., 2021).

Very few of these HOLC neighborhood descriptions explicitly mention crime. However, hundreds of neighborhoods achieving the lowest rankings ("definitely declining" and "hazardous") explicitly mention the presence of Italian immigrants as an adverse feature of the neighborhood. More over, one of the Mori immigrant enclaves in Cleveland, Ohio contains the following neighborhood description:

> This well-known Italian settlement has gained notoriety from the activities of the so-called "Mayfield Road mob" which figured prominently in gangster activities during prohibition and is said to be still active (Nelson et al., 2023).

While there is not enough information regarding concerns over criminality in the HOLC assessments to pursue the question empirically, we are able to test whether neighborhoods housing "Mori" immigrant enclaves were systematically "redlined" at differential rates. Here we specify a simple OLS regression, controlling for the neighborhood's Italian-born population proportion and the presence of any Sicilian born immigrants to see whether these "Mori" neighborhoods were more likely to achieve the lowest grade in HOLC assessments. The results from this regression are displayed in Table 2.

Even after controlling for the presence of Sicilians and the proportion of a neighborhood that is Italian-born, neighborhoods that housed immigrants from "Mori

Dependent variable	*HOLC Map Grade D ("Redlined")*			
	Model 1	Model 2	Model 3	Model 4
Mori	0.338***	0.293**	0.184**	0.168**
	(0.062)	(0.069)	(0.054)	(0.058)
Italian Population		0.376**		0.184**
		(0.0849)		(0.075)
Sicilians Present			0.227**	0.217**
			(0.052)	(0.054)
City FE	✓	✓	✓	✓
R-squared	0.144	0.148	0.169	0.170
DV Mean	0.392			
Observations	4,946			

Note: Standard errors are reported in parentheses and are clustered at the city level. Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 2: Table depicting OLS results of regressions of HOLC

commune" in 1920 faced a more than 40% higher risk of being redlined 15 years later. Aaronson et al. (2021) discusses some of the profound long run impacts that redlining, cutting neighborhoods off from formal lending, had on urban neighborhoods across subsequent decades. In the next sections, we will document that despite being "redlined" at higher rates than other comparable neighborhoods, these "Mori" enclaves were able to excel economically in the long run- a success that we attribute in part to the positive economic spillover effects of mafia presence.

3.3 Long term outcomes

Here, we document long term impacts of the mafia's presence in urban neighborhoods and showcase how these neighborhoods bucked the adverse trends experienced by most "redlined" neighborhoods. In doing so, we showcase results largely similar to those of Murphy & Rossi (2020), albeit in a vastly different setting- namely that the mafia wrought real and tangible long term benefits to the communities that it was embedded in.

To address this question, we regress long term outcomes, drawn from Opportunity Insights data (Chetty et al., 2018) on an indicator for 1920 "Mori commune"

immigrant presence. The results are broadly similar when instead adopting an instrumental variables approach, instrumenting 1959 mafia activity locales with the "Mori immigrant" instrument. Figure 4 graphically depicts some of these reduced form results. Models 2 and 3 from that figure show results with the inclusion of covariates to control for a neighborhood's 1920 Italian-immigrant proportion and the presence of Sicilian immigrants.

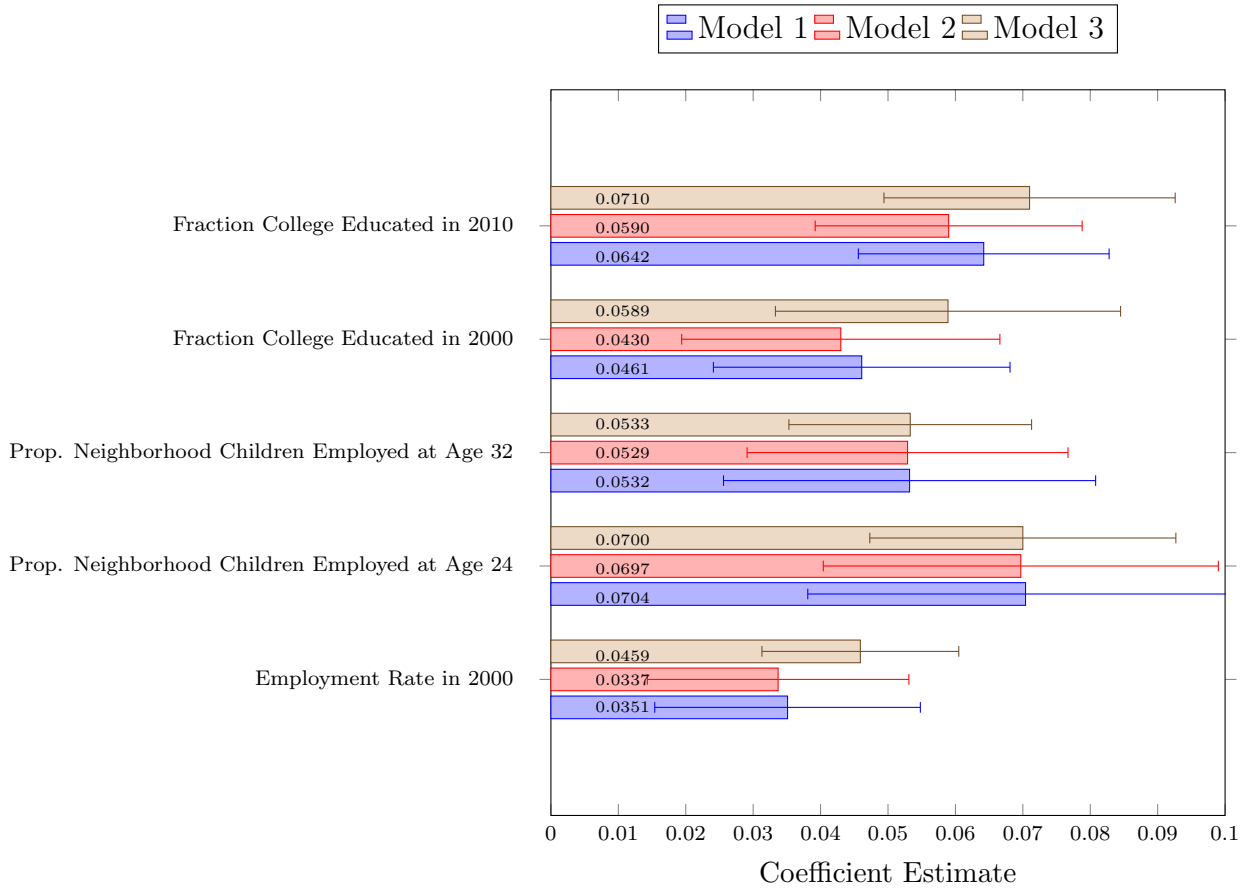


Figure 4: Figure depicting long term estimated long term impacts of mafia presence on neighborhoods with confidence intervals.

These results demonstrate that even when compared to similarly Italian 1920 neighborhoods that also housed Sicilian immigrants, better outcomes were achieved decades later in the areas where the mafia established roots. Namely, employment rates in 2000 were higher, more of the neighborhood had achieved college degrees, and children born in the neighborhood between 1978 and 1983- 50 years after Mori's

expulsion of the mafia from Sicily- were much more likely to be employed in young adulthood. These, along with other long term outcomes, showcase the long term impacts brought by the influx of an alternative form of capital into these neighborhoods and mirror the results shown for cartel villages in Mexico in Murphy & Rossi (2020).

4 References

Abramitzky, R., Boustan, L. P., Jácome, E., Pérez, S., & Torres, J. D. (2023). Law-Abiding Immigrants: The Incarceration Gap Between Immigrants and the US-born, 1870–2020. American Economic Review: Insights.

Kasey Buckles, Adrian Haws, Joseph Price, and Haley Wilbert. "Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project." Working Paper, 2024.

Aaronson, D., Hartley, D., & Mazumder, B. (2021). The effects of the 1930s HOLC "redlining" maps. American Economic Journal: Economic Policy, 13(4), 355-392.

Card, D. "Immigration and Inequality." American Economic Review, 2009.

Di Poppa, G. "The Long-term Effects of the Expulsion of Criminals." Working Paper, 2024.

Fishback, P. V., LaVoice, J., Shertzer, A., & Walsh, R. P. (2023). The HOLC maps: How race and poverty influenced real estate professionals' evaluation of lending risk in the 1930s. The Journal of Economic History, 83(4), 1019-1056.

Light, Jennifer. 2010. "Nationality and Neighborhood Risk at the Origins of FHA Underwriting." Journal of Urban History 36 (5): 634–71.

Murphy, T. E., & Rossi, M. A. (2020). Following the poppy trail: Origins and consequences of Mexican drug cartels. Journal of Development Economics, 143, 102433.

Nelson, R. K., Winling, L, et al. (2023). Mapping Inequality: Redlining in New Deal America. Digital Scholarship Lab. https://dsl.richmond.edu/panorama/redlining.

Sviatschi, M. M. "Making a Gangster: Exporting US Criminal Capital to El Salvador." Working Paper, 2022.