BROWN
Orlando Bravo Center
for Economic Research

Competition in the Black Market: Estimating the Causal Effect of Gangs in Chicago†

Bravo Working Paper # 2021-004

Jesse Bruhn †

Abstract: I study criminal street gangs using new data that describes the geospatial distribution of gang territory in Chicago and its evolution over a 15-year period. Using an event study design, I show that city blocks entered by gangs experience sharp increases in the number of reported batteries (6%), narcotics violations (18.5%), weapons violations (9.8%), incidents of prostitution (51.9%), and criminal trespassing (19.6%). I also find a sharp reduction in the number of reported robberies (-8%). The findings cannot be explained by pre-existing trends in crime, changes in police surveillance, crime displacement, exposure to public housing demolitions, reporting effects, or demographic trends. Taken together, the evidence suggests that gangs cause small increases in violence in highly localized areas as a result of conflict over illegal markets. I also find evidence that gangs cause reductions in median property values (-$8,436.9) and household income (-$1,866.8). Motivated by these findings, I explore the relationship between the industrial organization of the black market and the supply of criminal activity. I find that gangs that are more internally fractured or operate in more competitive environments tend to generate more crime. This finding is inconsistent with simple, market-based models of criminal behavior, suggesting an important role for behavioral factors and social interactions in the production of gang violence.

Keywords: Competition, Crime, Gangs, Neighborhoods, Urban, Violence
JEL classification: J46, K24, O17, R23, Z13

---

† Previously circulated under the title "The Geography of Gang Violence."

* Economics Department, Brown University. jesse_bruhn@brown.edu. / jessebruhn.com

I am indebted to Kevin Lang, John Friedman, Daniele Paserman, Ray Fisman, Alex Mas, Sam Bazzi, Pascual Restrepo, Phillip Ross, Bruno Martins, Arthur Smith, Felix Koenig, Ellora Derenoncourt, Andrew Jordan, Stephen Billings, Matt Pecenco, David Skarbek, and Bryce Steinberg for helpful conversations and comments. I would also like to thank Patrick Burke for introducing me to gang members and CPD officers in Chicago, Zach Jones for talking to me about his work collecting gang folklore, and Mick Dumke for providing me with advice about the Chicago FOIA process. I would like to thank the law firm Loevy and Loevy for filing a lawsuit against the CPD on my behalf. I would also like to thank Sarah Daly and Elena Barham for providing me with the CPD gang map from 2018. Finally, I would like to thank Jake Fabian for providing expert research assistance.

Competition in the Black Market: Estimating the Causal Effect of Gangs in Chicago[†]

Jesse Bruhn[⋆]

First version: November 1, 2019

This version: April 30, 2021

Abstract

I study criminal street gangs using new data that describes the geospatial distribution of gang territory in Chicago and its evolution over a 15-year period. Using an event study design, I show that city blocks entered by gangs experience sharp increases in the number of reported batteries (6%), narcotics violations (18.5%), weapons violations (9.8%), incidents of prostitution (51.9%), and criminal trespassing (19.6%). I also find a sharp reduction in the number of reported robberies (-8%). The findings cannot be explained by pre-existing trends in crime, changes in police surveillance, crime displacement, exposure to public housing demolitions, reporting effects, or demographic trends. Taken together, the evidence suggests that gangs cause small increases in violence in highly localized areas as a result of conflict over illegal markets. I also find evidence that gangs cause reductions in median property values (-$8,436.9) and household income (-$1,866.8). Motivated by these findings, I explore the relationship between the industrial organization of the black market and the supply of criminal activity. I find that gangs that are more internally fractured or operate in more competitive environments tend to generate more crime. This finding is inconsistent with simple, market-based models of criminal behavior, suggesting an important role for behavioral factors and social interactions in the production of gang violence.

Keywords: Competition, Crime, Gangs, Neighborhoods, Urban, Violence

JEL classification: J46, K24, O17, R23, Z13

[†] Previously circulated under the title "The Geography of Gang Violence."

[⋆] Economics Department, Brown University. jesse_bruhn@brown.edu. / jessebruhn.com

I am indebted to Kevin Lang, John Friedman, Daniele Paserman, Ray Fisman, Alex Mas, Sam Bazzi, Pascual Restrepo, Phillip Ross, Bruno Martins, Arthur Smith, Felix Koenig, Ellora Derenoncourt, Andrew Jordan, Stephen Billings, Matt Pecenco, David Skarbek, and Bryce Steinberg for helpful conversations and comments. I would also like to thank Patrick Burke for introducing me to gang members and CPD officers in Chicago, Zach Jones for talking to me about his work collecting gang folklore, and Mick Dumke for providing me with advice about the Chicago FOIA process. I would like to thank the law firm Loevy and Loevy for filing a lawsuit against the CPD on my behalf. I would also like to thank Sarah Daly and Elena Barham for providing me with the CPD gang map from 2018. Finally, I would like to thank Jake Fabian for providing expert research assistance.

1 Introduction

Gangs[1] are an integral component of the supply chain for illegal drugs. As the end point of sale for illicit narcotics, gangs represent the consumer-facing side of a global industry that the United Nations estimates is worth over $420 billion[2] annually (United Nations Office on Drugs and Crime, 2005), yet direct evidence on the causal effect of gangs is scarce. This is largely because reliable data on street-level criminal enterprise are rare. Further complicating matters is that gangs tend to operate in lower socioeconomic status neighborhoods (Levitt and Venkatesh, 2000, 2001). Thus even when it is possible to observe gang activity, any direct comparisons across gang-occupied and unoccupied areas will be contaminated by selection.

In this paper, I study criminal street gangs using new data from Chicago that describes the geospatial distribution of gang territory and its evolution over a 15-year period. This is the first and only administrative panel data set of its kind in the literature,[3] and it allows me to address key identification challenges concerning the selection of gangs into violent neighborhoods. The data come from the Chicago Police Intelligence division and are based on the firsthand accounts of officers, confidential informants, and police administrative records (City of Chicago Office of the Inspector General, 2019). I obtained the data via a nearly two-year-long Freedom of Information Act (FOIA) process that involved multiple requests and appeals. After obtaining the maps, I took steps to qualitatively validate their accuracy by visiting illegal drug markets in Chicago where I interviewed gang members, community residents, and police officers.

The first contribution of this paper is to document a new set of stylized facts about the distribution and dynamics of gang territory. I begin by showing that Chicago has a large number of small gangs with fluid boundaries. In an average year, there are 57 gangs in the data, and of those gangs an average of 29.6 have moving boundaries. I also find evidence

[1] While the formal definition of a gang varies widely across state and federal law (National Gang Center, 2019), the 2006 edition of the Chicago Crime Commission's "Gang Book" offers the following representative definition: "A street gang is an organized group that participates in criminal, threatening, or intimidating activity within the community. This anti-social group, usually of three or more individuals, evolves from within the community and has recognized leadership as well as a code of conduct. The group remains united during peaceful times as well as during times of conflict. A street gang is an organization that exhibits the following characteristics in varying degrees: 1. a gang name and recognizable symbols; 2. a definable hierarchy; 3. a geographic territory; 4. a regular meeting pattern; 5. a code of conduct; 6. an organized, continuous course of criminal activity" (Kirby et al., 2006).

[2] After adjusting for inflation

[3] Notably, Papachristos (2009) and Papachristos et al. (2013) pioneered the use of Geographical Information System (GIS) information from police intelligence to study gangs. These authors had access to a single map of gang-controlled territory in Chicago for an unspecified year. This is the first paper to have access to multiple shapefiles linked to different points in time.

of gang "fracturing." Over this period, the number of distinct groups per member operating within the territory of a typical gang increased by twofold. This finding aligns with recent qualitative sociological work suggesting that most gangs in Chicago no longer operate as cohesive entities (Aspholm, 2020; Great Cities Institute, 2019). Next, I show striking patterns in the geospatial distribution of violence that are suggestive of gang conflict. Notably, I find that while there are large differences in violent crime across gang-occupied and unoccupied neighborhoods, these differences are well predicted by proximity to a potential rival gang. In particular, violence at the border of gang territory is incredibly volatile. The probability of observing a more than two standard deviation outlier in the distribution of violent crime is 107% higher in a border region than it is in other gang-controlled areas.

Finally, I show that the distribution of gang territory largely mirrors racial and ethnic cleavages within the city. For example, territory controlled by gangs identified as black contains a population that is 85% black and 7% Hispanic, while territory occupied by Hispanic gangs is 8% black and 64% Hispanic. This finding suggests that gangs may find it more difficult to occupy neighborhoods that differ from them demographically. Consistent with this fact, I use historical data on gang rivalries in Chicago to document that gang conflict typically occurs within race and ethnicity and rarely across.

This paper's primary contribution is to provide evidence that gangs cause small increases in violence in highly localized areas as the result of conflict over illegal markets. Using an event study design, I show that city blocks that are entered by gangs experience sharp increases in reported batteries (6%), narcotics violations (18.5%), incidents of prostitution (51.9%), weapons violations (9.8%), and criminal trespassing (19.6%). Within these broad categories, I find that the specific types of criminal acts that shift upon gang entry are consistent with qualitative accounts of how gangs operate illegal markets. For example, the increase in narcotics is disproportionately driven by the possession and sale of crack cocaine, and the increase in weapons violations is disproportionately driven by incidents involving possession, sale, or use of a handgun.

Interestingly, I find that gang entry is also accompanied by a sharp reduction in the number of reported robberies (-8%), consistent with the idea that gangs use violence to exert control over their neighborhood as a means of protecting the integrity of the market. Importantly, I find that gangs do not cause increases in crimes that have no obvious connection to gang activity such as kidnapping, fraud, or electronic harassment. I also show that the findings cannot be explained by pre-existing trends in crime, changes in police surveillance, complicated geospatial trends in policing practice, crime displacement, exposure to public housing demolitions, reporting effects, or demographic trends. Taken together, the evidence suggests that these estimates do, in fact, represent the causal effect of gangs on the neighborhoods they inhabit. I also find evidence that gangs cause reductions in median property values (-$8,436.9) and household income (-$1,866.8), suggesting the pernicious effects of gangs extend beyond the domain of crime and into other neighborhood-level outcomes.

As the final contribution of this paper, I examine whether the causal effect of gang entry varies with the illegal market's industrial organization. Ever since Becker (1968), economists have sought to understand whether and to what degree standard models of economic behavior describe agents in the black market. Nobel laureate James Buchanan famously argued that allowing criminal organizations to obtain market power could be socially beneficial (Buchanan, 1973): "If monopoly in the supply of 'goods' is socially undesirable, monopoly in the supply of 'bads' should be socially desirable." Empirically, there is a small but convincing literature that documents instances of criminal behavior responding rationally to market incentives (e.g., Olken and Barron, 2009; Arora, 2018; Brown et al., 2020).

By way of contrast, there is a recent and growing body of work that suggests crime is often the product of non-market forces. Famously, Glaeser et al. (1996) argued that the spatial variance of crime is too large to be the product of exogenous changes to the cost and benefit of particular criminal acts. As an alternative hypothesis, they suggest that social interactions may play an important role. This hypothesis has subsequently accumulated a small body of empirical support (e.g., Jacob and Lefgren, 2003; Durlauf and Tanaka, 2008; Bayer et al., 2009; Durlauf and Ioannides, 2010; Billings et al., 2019). Recent work has also emphasized the role of important behavioral factors and cognitive biases in the production of violence. For example, Heller et al. (2017) find that interventions that force at-risk youth in Chicago to reflect on automatic thought processes and behaviors can have transformative effects on their propensity to commit violent crimes.

Motivated by these different theories of criminal behavior, I explore the role of market competition in the production of crime. To fix ideas, I formalize Buchanan's theory via a stylized Cournot model of gang competition that features endogenous law enforcement. The model predicts that the causal effect of gang entry will be smaller in more competitive environments and the causal effect of gang entry will be smaller when the elasticity of law enforcement effort is high. I find the opposite pattern in the data. The causal effect of gang entry on narcotics is *larger* at gang borders and when a second gang enters the market. I find similar, but noisier, results for violence. I also find that gangs that have more internal competition as a result of fracturing generate larger causal effects on narcotics and violence than more cohesive gangs. This last finding is of independent interest, as it provides the first empirical support for sociological and criminological theories about the role of gang fracturing in the creation of violence (Aspholm, 2020; Great Cities Institute, 2019).

Next I show that the causal effect of gangs on crime does not appear to vary with distance to the nearest police station. Under the assumption that it is more costly for police officers to patrol more remote areas of the city, this finding provides evidence that the causal effect of gang entry on crime does not vary with the elasticity of police effort. Taken together, the evidence implies that the industrial organization of crime in Chicago is not well described by simple market-based models of criminal production. This suggests that social interactions and behavioral factors likely play an important role in explaining gang violence in Chicago.

The findings in this paper contribute to a larger policy discussion surrounding gangs, neighborhoods, and law enforcement. Gangs are a consequential and persistent feature of life in large urban areas. A 2012 survey of law enforcement agencies in the United States found that over 3,100 jurisdictions experienced gang problems, including virtually all agencies serving cities of greater than 100,000 people (National Gang Center, 2019). This paper quantifies the cost that such gangs impose on the neighborhoods they inhabit, which is an essential ingredient for any cost-benefit calculation that will determine the appropriate scale of the policy response. Further, the finding that gangs do not appear to respond rationally to market incentives is useful for tactical law enforcement decisions regarding when and where to concentrate anti-gang resources.

Chicago is an interesting and important case study for trying to understand broader issues related to gang violence. According to the National Gang Center, Chicago is one of two "gang capitals" in the United States (National Gang Center, 2019).[4] Further, Hagedorn and Macon (1998) argue that Chicago has been important for the historical development and dissemination of important aspects of gang norms and culture in the United States, particularly in the industrial Midwest. Thus there is reason to expect that lessons learned in Chicago will apply in the US more broadly.

More recently, Chicago saw decades of progress reverse in 2016 as the number of murders suddenly and inexplicably climbed to its highest level in almost 20 years (Sanburn and Johnson, 2017). A recent report from the University of Chicago Urban Labs noted that the violence in 2016 was "carried out by teens and young adults in public places," involved guns, and often stemmed from an altercation. It concluded that most candidate explanations[5] were not consistent with a straightforward examination of the data; however, limited information on gang activity left open the possibility that gangs were the root cause (Kapustin et al., 2017).

[4] The second is Los Angeles

[5] These included factors such as poverty, segregation, guns, demolition of public housing, changes in police behavior, changes in funding for social services, family structure, shuttering of mental health facilities, and weather.

Thus as a pure public policy concern, understanding gang violence in the third largest city in the United States is necessary and broadly useful for developing interventions that will push back against these negative trends.

2 Related literature

There is now a small body of compelling evidence exploring the role of gangs on economic development in Central and South America (e.g., Dell, 2015; Carvalho and Soares, 2016; Monteiro and Rocha, 2017; Sviatschi, 2019b,a). Most closely related to this project are two contemporaneous papers. Melnikov et al. (2019) leverages a spatial regression discontinuity design to estimate the causal effect of cartels on households in San Salvador, Ecuador. They find that earnings and labor mobility are lower among households just inside the cartel's territory than among those that are just outside. Sobrino (2019) uses an event study design and data on cartel activity from Google News events to explore the relationship between cartels and violence in Mexico.

However, there are crucial differences between gangs in Central and South America and those found in other parts of the world. For example, the cartels of Medellín, Colombia tend to be highly vertically integrated, with a well-defined organizational hierarchy among criminal enterprises that in turn take on state-like functions (Blattman et al., 2019, 2021). By way of contrast, gangs in cities like Chicago tend to be much smaller, less organized, and more fractured (Aspholm, 2020; Great Cities Institute, 2019). For these reasons, it is not clear whether this existing body of work provides useful guidance for policymakers and researchers seeking to understand the trade-off between the social cost and benefit of anti-gang efforts in the rest of the world.

There is also a small literature within economics that explores the nature of gang activity in the United States. Grogger (2002) demonstrates that gang injunctions reduced crime in Los Angeles. Cook et al. (2007) finds that the market for guns in Chicago is thin because gangs find them too risky to sell. Bruhn (2018) finds that public housing demolitions in Chicago increased crime city wide due to gang members being forced out of their traditionally defined territories. This literature is complemented by descriptive work at the intersection of economics and political science that analyzes the internal institutions of criminal organizations (e.g., Leeson, 2007; Skarbek, 2010; Dooley et al., 2014). Levitt and Venkatesh (2000, 2001) also provides classic descriptive work that analyzes the finances of a drug-selling gang in Chicago as well as the later life outcomes of some Chicago youth who were involved in gang activity.

A larger literature explores the nature of the European mafia. Several papers find that criminal institutions in Europe had important long-run consequences for local economic development (Pinotti, 2015a,b; Acemoglu et al., 2017; De Feo and De Luca, 2017; Lonsky, 2019; De Martiis, 2020). There is also important work with a contemporary focus that documents the negative influence of the mafia on political corruption and the day-to-day efficiency of governmental operations (Barone and Narciso, 2015; Daniele and Dipoppa, 2017; Marcolongo, 2020; Pinotti et al., 2020). And recent work by Calamunci (2021) finds that criminal organizations have important consequences for firm performance.

Relative to the existing work on criminal organizations in the Americas and Europe, my paper adds to these important literatures in three ways. First, I provide the only causal estimates of the impact of gangs that leverages administrative data with panel variation in gang presence. This allows me to make strong claims to causality and directly address some possible sources of selection concerning the selection of gangs into violent areas that have only been indirectly addressed in prior work. Second, I complement existing work on the causal effects of gangs by providing the first exploration of gang activity in a setting where gangs are fractured and only minimally vertically integrated, which is important for establishing a nuanced view of how organized crime affects outcomes in different setting. Third, I provide new evidence on the role of market competition in gang activity, including the first direct test of the Buchanan (1973) model of black market competition.

Finally, there is also a rich and important tradition of qualitative and analytical work on gangs found in sociology and criminology. This body of work began with the seminal contribution of Frederick Thrasher who, in 1913, explored the operations and motivations of over 1,300 street gangs in Chicago[6] (Thrasher, 2013). More recently, Papachristos (2009), Papachristos et al. (2013), and Wildeman and Roberto (2015) find evidence that gang conflict in Chicago is well predicted by social, institutional, and network structures. This literature emphasizes the use of violence as a means of protecting the symbolic value of place, neighborhood, and social status among gang members. In economic terminology, they claim that gang violence emerges from social interactions rather than pure instrumental motives. Lindquist and Zenou (2019) review the criminological and economic literature that applies social network analysis techniques to criminal behavior. Taniguchi et al. (2011) find that, even after adjusting for demographics, street corners used for drug dealing or that have multiple gangs contain higher crime counts. And Brantingham et al. (2012) show that crime is higher at points that are equidistant from street corners controlled by gangs. My paper builds on this work by providing causal evidence consistent with the dominant view in sociology and criminology

[6]Thrasher also mapped the territory controlled by gangs of various ethnicities as well as corners/meeting spaces where they could be found.

regarding the deep drivers of gang violence in the United States.

3 Measuring gang territory

The data for this project come from FOIA requests sent to the Chicago Police Department (CPD), the city of Chicago open data portal, the 2000 census, and the 2006 edition of the Chicago Crime Commission's "Gang Book" (Kirby et al., 2006). In this section, I briefly summarize the primary data sets used in this analysis. I also describe steps I have taken to qualitatively validate the data via firsthand interviews of gang members, neighborhood residents, and CPD officers. More detailed descriptions of the data, as well as additional details regarding the data cleaning process, are in Appendix A.

3.1 Gang map data

The CPD maintains detailed Geographical Information System (GIS) records of gang territory. A recent report from the City of Chicago's Office of Inspector General contains the most in-depth publicly available discussion regarding how these maps are generated (City of Chicago Office of the Inspector General, 2019). According to the report, the maps themselves are produced from information gathered during annual "gang audits." During an audit, the district will complete a questionnaire on local gang activity by drawing on sources such as "Gang Arrest Cards, information from CI's [*sic*], and subject matter knowledge from department members such as representatives from the Gang Investigations Division." One portion of this questionnaire concerns gang territorial boundaries. Once the questionnaire is complete, it is reviewed by the CPD's Deployment Operations Center to verify accuracy through sources such as "first hand interviews, Contact Cards/Investigatory Stop Reports, reports of shots fired, and incidents." The information from the audit is then stored in the CPD's "Caboodle" database. Papachristos (2009) argues that, to a first approximation, these boundaries represent a "locus of control," as perceived by the gang members themselves.

From December of 2016 to June of 2017, I sent a series of FOIA requests and appeals to the CPD. In these requests, I asked the CPD to provide me with any and all GIS data related to gang territorial boundaries for as far back in time as they had maintained records. Most of these requests were denied, and typical denials cited officer safety as the reason.[7] Some requests were "fulfilled" by sending me non-responsive documents.[8] The CPD eventually provided me

[7] This was even though the CPD had released one of these maps some ten years earlier for publication in the Chicago Crime Commission's "Gang Book" (Kirby et al., 2006).

[8] For example, in response to one request, the FOIA office sent me the names and pictures of suspected gang

with a collection of shapefiles that describe gang territorial boundaries annually from 2004 to 2017. I shared the map data with a second research team in spring of 2019, and I also advised them on the FOIA process with the CPD. This team then sent a subsequent FOIA request to the CPD that resulted in a corresponding shapefile from 2018, which they generously shared with me. As of the date of this writing, the CPD has refused to acknowledge or answer any clarifying questions I sent them via email regarding these maps and the data they contain. For more detail on the procedures used to clean the gang map data, see Appendix A.1.

Gang member, neighborhood resident, and CPD officer interviews

To verify the accuracy of the gang maps, I solicited the help of a freelance journalist in Chicago who had sources within several of the gangs in my data as well as sources within the CPD. He invited me to meet with these sources and ask them questions, and in November of 2019, I spent four days in Chicago's West Side traveling through the Austin and North Lawndale neighborhoods to interview current and former gang members, neighborhood residents, and CPD officers.

I found that gang borders are well known to most gang members and neighborhood residents. Many of the individuals I interviewed were initially skeptical that the CPD would be able to accurately map gang territory. This appeared to stem both from the negative publicity that CPD's gang database has received in recent years[9] and from the fact that the individuals I interviewed in these communities did not feel a great deal of trust for the police department. For example, one former gang member I spoke with stated he did not believe anything that the CPD produced because they regularly harassed him on the basis of his race. However, after discussing specifics of the maps with them, virtually all of these individuals reported that the data for their local neighborhood was accurate.

The fact that the territories are so well known is likely because the gang borders are not secret. I observed several gangs openly selling drugs on street corners, abandoned property, and in the parking lots of convenience stores and gas stations. Further, neighborhood residents that we approached to interview could often describe gang boundaries in their neighborhood from memory. The fact that gang boundaries are so well known may seem surprising. However, one of the former gang members I interviewed said their physical safety was tied to knowing these boundaries. Thus the gang members (and likely the neighborhood residents as well) have a strong incentive to learn them.

---

members in each community area of the city. Since these were clearly not the maps I had asked for, I continued to file appeals.

[9] For example, see https://www.propublica.org/article/politic-il-insider-chicago-gang-database.

I also spent one evening in an open air drug market contained in a gas station south of Interstate 290.[10] While at the drug market, the journalist facilitated a discussion between me and several of the gang members as they sold heroin to their customers. During this conversation, I observed a second gang selling drugs out of the parking lot of the neighboring convenience store. After completing the interviews, the journalist and I reviewed the most recent CPD gang map for that region. We found that the border between these two gangs was accurately reflected in the map.

I also interviewed a CPD beat officer who was patrolling this neighborhood in his police vehicle. He showed us the computer system in his patrol car that would allow him to view the maps. When I asked him if he used the maps for his day-to-day police work, he claimed that he did not need to, because he already knew where the gangs were operating as a result of his firsthand observations and conversations with the neighborhood residents.

Finally, the journalist introduced me to a former high-ranking member of CPD's "gang unit" that is responsible for the production of the maps. Unfortunately, this officer had not been directly involved in the unit since the late 90s, so it is not clear whether the information he provided me is completely up to date. One concern I had was that the gang unit might not take the production of the maps very seriously. If the officers updated maps reluctantly or without accurately reviewing the information submitted as part of the district-level gang audits, this could result in substantial amounts of measurement error in the data. The officer assured me that this was not the case. He claimed that the gang unit believes these maps are essential when officers are responding to dangerous situations in sectors that they are unfamiliar with. In that case, knowing the relevant gang boundaries and actors in the area could prove important for officer safety and the conduct of quality police work. According to him, the CPD takes the production of these maps very seriously.

3.2 Additional sources of gang data

I supplement the gang map data with several other sources of information on Chicago gangs. First, I collected historical data about the gangs themselves from the 2006 edition of the Chicago Crime Commission's "Gang Book" (Kirby et al., 2006). The Chicago Crime Commission is a non-profit organization founded in 1919 as a partnership between law enforcement and the business community of Chicago (Chicago Crime Commission, 2019). The book is irregularly produced as an informational tool meant to help law enforcement agencies (in-

[10] This region of Chicago has been labeled by the media as the "heroin highway." For example, see https://www.npr.org/sections/thetwo-way/2014/06/13/321692592/chicago-heroin-highway-bust-shows-a-new-face-of-organized-crime.

cluding those outside of Chicago) better police gangs. This book is based on Chicago Crime Commission interviews with CPD officers as well as current and former gang members. It contains short narrative histories along with associated races and ethnicities for most of the gangs that appear in my data. I transcribed these races and ethnicities by hand and also transcribed information from the narrative histories about which gangs were historically rivals or at war with one another prior to 2006.[11] See Appendix A.3 for more detail.

I also make use of two additional elements from CPD's gang database. The first data element is known as the "strategic subject list" and was released by the CPD after the Chicago Sun Times filed a lawsuit against the department.[12] According to the Chicago open data portal, these data "represents a de-identified listing of arrest data from August 1, 2012 to July 31, 2016," which was used to rank individuals on their propensity to be involved in gun violence. The data contain a field that identifies an individual as a suspected gang member, allowing me to provide summary statistics related to the characteristics of these suspected gang members relative to the universe of arrested individuals.

The second data element contains sparse information on the universe of individuals that have been added to CPD's gang database. It was obtained by ProPublica as part of their investigative reporting and consists of two spreadsheets that cannot be linked.[13] The first spreadsheet contains data on each individual added to the database, the name of the gang to which they are suspected of belonging, and the specific gang "set"[14] to which the individual belongs. This spreadsheet allows me to estimate the number of individuals and the number of gang sets within each gang by birth cohort. The second spreadsheet contains fields related to age, gang, the date the individual was added to the database, and the police beat that added them to the database. This allows me to construct a police-beat-level proxy for police monitoring activity by measuring the number of individuals added to the database in each city police beat during each calendar year.

[11] The "Gang Book" also contains data on alliances, but outside of a gang's historical affiliation with the People Nation or Folk Nation, this information is much more irregularly recorded.

[12] See the following news article for more information about the lawsuit: https://chicago.suntimes.com/city-hall/2020/1/27/21084030/chicago-police-strategic-subject-list-party-to-violence-inspector-general-joe-ferguson.

[13] See the following website for more information about the data: https://www.propublica.org/datastore/dataset/chicago-police-clear-gang-data

[14] A gang "set" or crew is an independent group that operates in a semi-independent manner within a larger gang. See Aspholm (2020) for more detail.

3.3 Crime data, census data, and aggregation procedures

From the Chicago open data portal, I downloaded incident-level crime data spanning the years 2001 to 2018 along with shapefiles containing census block and police beat boundaries. As discussed in more detail in Herrnstadt et al. (2018), the incident-level crime data contain the universe of reported crime in Chicago over this period. For each crime, the data documents the associated Illinois Uniform Crime Reporting (IUCR) code, the date and time when the crime was reported to have occurred, along with grid coordinates where the crime was reported to have happened. I use the census block shapefiles to aggregate incidents to counts at the census block-year-IUCR level.[15] The crime data contain a "description" field with 528 distinct codes that provide additional detail about the nature of the crime. I also generate an alternative aggregation based on these codes that I use for heterogeneity. See Appendix A.2 for more detail on this aggregation.

I measure the position of each of these census blocks relative to gang territory by calculating, for each year in the gang map data, the minimum distance of each census block centroid to the boundary of each gang polygon. In all cases, I denote centroids contained within a gang polygon with a negative distance to the gang border; centroids that are outside the boundaries of a given gang are denoted with a positive distance. I supplement the crime data and gang map data with block-level population counts by race, ethnicity, and age from the year 2000 decennial census. I also measure some additional time-varying outcomes, using American Community Survey (ACS) five-year averages at the census tract level. I drop 190 blocks associated with the Chicago airport as well 9 blocks that are contained in the Chicago open data portal shapefiles but are not included in the census data for Cook County, Illinois. After these restrictions, I am left with a balanced panel of 24,490 city blocks observed over 15 years.

4 Summary statistics and stylized facts

I begin with a description of the gang members themselves, which is useful for two reasons. First, it provides general background knowledge necessary for interpreting the causal effect of gang entry. In the main analysis, I treat a gang as synonymous with its territory. Behind that territory are the individuals who actually act as members of the gang, and this is the only descriptive evidence available regarding their characteristics. Second, observing the crime profiles of arrested gang members helps to calibrate expectations regarding what types of

[15] The grid coordinates in the CPD crime data contain a small amount of spatial noise; however, the documentation notes that this noise should not change the block where the crime was reported to have occurred.

crimes we might expect to see gangs influence when we arrive at the causal analysis. It would be concerning if gang members are typically involved in drug crimes, but the causal analysis detected effects on traffic violations. Thus Table 1 presents summary statistics for suspected gang members and non-gang members from the universe of individuals over 18 who were arrested in Chicago between the years of 2012 and 2016.

[Table 1 about here.]

Gang members are young, male, and disproportionately likely to be a victim/suspect of a violent crime. There is an 8 percentage point difference between suspected gang members and non-gang members in the probability that they are between the ages of 20 and 30 years old. While the differences in the less than 20 age group is smaller than this gap, this is almost certainly a lower bound since the data are censored below age 18. Gang members are also 25 percentage points more likely to be male.

The differences in crime victimization rates across group are staggering. Gang members are ten times as likely to be the victim of a shooting and are four times as likely to be a victim of a battery. Conversely, gang members are roughly nine times as likely to be arrested for using a weapon and are four times as likely to be arrested for violence. The fact that the victimization rates mirror the rates at which gang members commit crime suggests that much of the bloodshed may be directed at one another rather than the community at large.

Gang members are also disproportionately likely to be arrested for narcotics violations. While gang members are roughly twice as likely to have been arrested for exactly one narcotics violation, they are seven times as likely to have been arrested for multiple drug offenses. This finding highlights the primary role that the drug trade plays among street gangs.

Next I demonstrate that there are a large number of gangs in Chicago and most of them are small. Figure 1 displays estimated membership for every gang found in the map data in 2004. To estimate membership, I use the anonymized CPD gang database to count the number of individuals within each gang who belong to the 1975–1990 birth cohorts. The logic behind using these birth cohorts is that these individuals would range in age from 14 to 29 as of the year 2004. However, the diagram is very similar and yields the same broad conclusions if I measure gang "size" using the territory area instead of membership. See Appendix B.1.

[Figure 1 about here.]

With that in mind, it is important to note that the gangs labeled in these maps do not necessarily represent cohesive units. As described in Aspholm (2020), these gangs are themselves composed of independent groups known as gang sets that will often "clique up" or form alliances with other sets from different gangs. Further, the qualitative work suggests that the importance of these gang sets has increased over this period as gangs have become more fractured.

In Figure 2, I document this fracturing by plotting the number of gang sets per member by birth cohort for the five largest gangs in the data. In line with the qualitative work, I find that the prevalence of these gang sets begins to sharply increase with the 1993 birth cohort. This cohort is interesting because they would have begun to "age in" to crime shortly after the demolition of Chicago's most notorious high rise housing projects, which some scholars have hypothesized is responsible for the increase in gang fracturing.[16] This is the first quantitative evidence that gang fracturing has increased over time.

[Figure 2 about here.]

[Figure 3 about here.]

Geographically, the gangs are distributed throughout most of the city. Figure 3 plots gang-controlled area in the years 2004 and 2018. To facilitate visualization, in Figure 3 I have highlighted the five largest gangs by area in 2004 and collapsed the territory of the remaining gangs into a sixth polygon. In an average year, there are 57 gangs present in the data, and they occupy 32% of the city's land mass. In 2004, the area occupied by gangs accounted for approximately 62% of the city's population as measured in the year 2000 census.

Observe that gang territory is highly non-convex. For example, in 2004 the Vice Lords primarily controlled small patches of territory dotted along the West Side of Chicago while also maintaining a small presence deep in the South Side. These southern patches of Vice Lord territory are surrounded on all sides by Gangster Disciple and Black P. Stones neighborhoods.

Gang territory also frequently shifts over time. The total fraction of the city under gang control ranges from 28% in 2012 to 38% in 2009. Most gang boundaries (86.1%) move at least once over the course of the sample frame, and in an average year, 29.6 gangs have boundaries that move.

Gang territory is also violent. Figure 4 plots reported batteries at the block level in 2004 against distance to the nearest gang border. The figure also includes conditional means of crime within 70 equal-length distance bins. I have reduced the opacity of individual census

[16] Aspholm (2020) provides a qualitative account, and Bruhn (2019) provides quantitative evidence on the impact of the demolitions on gang conflict. While tempting, these demolitions are unlikely to prove useful as a source of identifying variation in the empirical section of this paper. This is because the demolitions also triggered large amounts of within-city migration and were often accompanied by neighborhood revitalization efforts, both of which render the exclusion restriction untenable.

block observations so that the darkness of the dots can be taken as a rough visual representation of the density. Negative values of distance denote blocks that are inside gang territory; positive values denote blocks that are outside gang territory. Note that the figure looks similar if adjusted for population by replacing crime counts with residuals from a regression of violence on year 2000 population on the y-axis.[17] Thus, the patterns in this figure are unlikely to be entirely the result of population differences between occupied and unoccupied territory (see Appendix B.2 for more detail).

[Figure 4 about here.]

There are two patterns of note in this figure. First, observe that the cross-sectional variance of violence is large ($\sigma = 6.2$ relative to a mean of 3.5). Also, the most extreme observations are clustered almost exclusively in a narrow region approximately 0.25 miles wide around the border. The probability of observing a more than two standard deviation outlier in the distribution of violent crime is 107% higher in this border region than it is in other gang-controlled areas. Second, the conditional means suggest that there is a discrete change in violence that occurs at the gang boundary. The fact that this is visible despite the y-axis being scaled to accommodate the extreme outliers at the border suggests that this shift is a quantitatively important feature of the data-generating process. In fact, a simple cross-sectional regression of violent crime in 2004 on a gang occupation indicator[18] with no other controls can explain 8.7% of the variation[19] in the data. Is this level shift the result of selection or is it the causal effect of gangs? I provide direct evidence on causality in Section 6, but for now I present descriptive evidence consistent with the idea that gang conflict plays an important role.

Figure 5 plots reported batteries against distance to a potential rival gang. Each point in this figure is a census block observed in 2004. The subpanels effectively hold distance to nearest gang constant by restricting the sample to blocks whose centroids are within a specified distance of the nearest gang border. Thus each panel of this figure approximates an isoplane from the three-dimensional surface mapping violence into distance to first and second closest gang.

[Figure 5 about here.]

[17] There are a non-trivial number of census blocks in the city with zero population that also frequently experience crime; hence I prefer to residualize rather than scale the dependent variable by year 2000 population levels. However, the results are similar if I drop census blocks with zero population and put batteries per capita on the y-axis instead. See Appendix B.2 for more detail.

[18] Defined such that the indicator takes a value of one if the block's centroid is within the territorial boundary of any gang.

[19] As measured via adjusted R-squared

Observe that the mean and variance shifts discussed in relation to Figure 4 appear to be almost entirely driven by blocks that are close to the border of two gangs. Within gang territory, blocks that are further than a half mile from a potential rival do not appear to be systematically different than blocks that are over a half mile from any gang.

Finally, the distribution of gang territory in Chicago mirrors the racial and ethnic cleavages of the city. Figure 6 describes the racial/ethnic composition of blocks occupied by gangs identified with a specific race/ethnicity. A block is considered occupied if its centroid falls within a gang's territory. Territory is measured using the 2004 gang maps, and racial composition is measured via the 2000 census. Each subpanel of Figure 6 corresponds to territory occupied by gangs identified in Kirby et al. (2006) as having the indicated primary racial/ethnic identity. Bars are shaded according to the race/ethnicity categories denoted on the y-axis to facilitate visual comparisons across subfigures. Note that because individuals in the census may identify with both a race and ethnicity (e.g., black and white Hispanic), the bars do not sum to one.

[Figure 6 about here.]

Figure 6 establishes that race is an important predictor of where gangs operate in the city. In fact, territory controlled by gangs identified as black contains a population that is 85% black and 7% Hispanic, while territory occupied by Hispanic gangs is 8% black and 64% Hispanic. This pattern suggests that gangs may find it more difficult to occupy areas far from where they live or that otherwise mismatch with their racial identity. Thus if the sociological literature is correct, and gang conflict is driven by the desire to exert influence over neighborhoods,[20] we would expect the majority of fighting to be within rather than across race.

Figure 7 displays the historical network of gang conflict. Each node in this diagram represents a gang from the CPD data. The size of the nodes corresponds to gang membership as measured via the CPD gang database over the 1975–1990 birth cohorts. The shading of the nodes denotes the gang's primary race/ethnicity as described in Kirby et al. (2006). Edges signify that the narrative history contained in Kirby et al. (2006) described the two connected gangs as being in conflict at some point in the past.

[Figure 7 about here.]

From the figure, two important patterns emerge. First, historically, gang conflict occurred within rather than across race. Second, within racial subnetworks, conflict appears to be

[20] As discussed in Papachristos (2009) and Papachristos et al. (2013)

organized around a central, large gang, perhaps because large gangs are more likely to bump borders with potential rivals than small gangs. In any case, these facts are consistent with the idea that gangs fight over neighborhoods in the presence of frictions emerging from racial mismatch.

5 Identification and estimation

Defining the causal effect of interest in this context is not straightforward. For example, one could imagine defining the causal effect of interest in terms of a large-scale experiment that takes every block in the city and randomly allocates it to the "locus of control" of a gang[21] or into an unoccupied control group. The problem with this thought experiment is that such manipulation would break important links between the gangs and their neighborhood of origin. These links have been stressed by the sociological and qualitative literature[22] and appear to be empirically relevant given the striking patterns related to race and conflict found in Section 4. It seems unlikely that this experiment would uncover anything like the causal effect of Chicago gangs conditional on the territorial distribution we observe today.[23]

For this reason, I define the causal effect of interest in terms of a more limited experiment. Thus in the ideal experiment, I would take city blocks that have some likelihood of gang occupation and randomly add some of them to the territory of the nearest gang. This experiment would identify the causal effect of occupation by the average gang[24] on Chicago neighborhoods with some positive probability of becoming occupied. This is the policy-relevant treatment effect in this context.

However, as observed in Section 4, gang territory is not randomly allocated throughout the city. Hence any observational comparison between gang-occupied and unoccupied areas is likely to be contaminated by selection. It is also possible that violence could cause gang occupation and/or formation. To address this and other sources of selection, I leverage the panel variation in the data via a matched event study design.

[21]Perhaps by randomly providing the gangs with high-powered incentives to occupy certain street corners or neighborhoods

[22]For examples, see Aspholm (2020); Papachristos (2009); Papachristos et al. (2013).

[23]Another way to see the problem with this hypothetical experiment is to consider that the University of Chicago has its own dedicated police force, and I think most economists who have spent time in Hyde Park over the last 20 years would agree that there is no universe over this time period where a gang was going to open a drug market outside of Saieh Hall. Thus randomly allocating the census block that contains Saieh Hall to gang territory does not seem like an empirically relevant experiment.

[24]Approximately weighted by the size of the pre-existing distribution of each gang's territory

5.1 Matched event study design

I define a gang occupation event as occurring when a census block has a gang within a quarter mile of its centroid at time t but did not have one at time $t-1$. I chose the quarter mile threshold because that is approximately the distance where the mean shift occurs in crime outcomes between occupied and unoccupied areas as documented Section 4; however, in Section 7 I allow the treatment to vary more flexibly with distance to a gang. I then use all of the crime variables I can observe in 2001 and 2002, along with demographic and housing stock variables from the year 2000 census, to estimate the probability of a gang occupation event. From there, I propensity score match each block that experiences an occupation event at some point during the sample frame to one that does not. I also require that the matched control blocks do not reside in the same police district as the treated block, to ensure that the comparison group is not contaminated by the gang occupation event. However, the results are not sensitive to details of the matching procedure (see Appendix D.1). For more detail on the matching procedure itself and summary statistics related to common support and balance, see Appendix C.

From there, I estimate variations of the following model:

$$y_{bt} = \delta_b + \delta_{mt} + \sum_{\tau \in T} \beta_\tau d_{b\tau} + \epsilon_{bt}, \tag{1}$$

where y_{bt} is crime in block b at time t, δ_b is a block fixed effect, $m = m(b, b')$ is the mapping between blocks and their associated match pair index (m), δ_{mt} are time-by-match pair fixed effects, and $\tau = t - t_m$ denotes time relative to the period (t_m) that the occupied block in match pair m becomes occupied. $T = \{-5, -2\} \bigcup \{0, 5\}$ is the set of event times, $d_{b\tau}$ is an indicator that takes a value of one when block b is occupied at event time τ, and ϵ_{bt} is a residual.

The coefficients of interest in regression (1) are β_τ. Note that I balance the model in event time[25] and except for the period prior to treatment (event time $\tau = -1$), I saturate the model in event time indicators. This makes the coefficients of interest here equivalent to estimating the corresponding 2x2 difference-in-difference design that would identify each β_τ separately by match pair, then averaging these 2x2 dif-in-dif coefficients across match pairs. This straightforward interpretation is desirable because it allows me to avoid complications related to the

[25]This is approximately true. In all regressions, I restrict the sample to a range of event-times such that at least 90% of the match pairs will contribute identifying variation. This allows me to observe a relatively lengthy period of time before and after treatment while also ensuring that the event study coefficients are not influenced by the changing composition of the sample used to identify them. Because crime from the years 2001 and 2002 were used in the matching step, I do not include these years when estimating the event study.

OLS conditional variance weights that emerge in two-way fixed effect models that do not satisfy these criterion (see Callaway and Sant'Anna, 2018; de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2019; Miller et al., 2019; Roth and Sant'anna, 2021; Sun and Abraham, 2020, for examples of how alternative modeling choices can generate misleading results).

Thus, the coefficients of interest in this design represent the causal effect of gang occupation at event time τ provided the trend for the matched comparison blocks is (on average) an accurate counterfactual for the trend the occupied would have otherwise experienced. This parallel trend assumption is formalized mathematically as $\mathbb{E}(\epsilon_{bt} d_{b\tau}) = 0$.

To summarize the results in terms of an average effect, I also estimate models of the form

$$y_{bt} = \delta_b + \delta_{mt} + \beta d_{bt} + \epsilon_{bt}, \tag{2}$$

where d_{bt} is an indicator that takes a value of one in all periods after gang occupation and is zero otherwise and the remaining variables/parameters are as previously defined. The coefficient on the pre-post indicator here (β) is analogous to a classic dif-in-dif in that it is estimated by comparing the difference between all of the post-periods and all of the pre-periods across treatment and control groups within each match pair and thus identifies a causal effect under a similar parallel trend assumption as in equation (1).[26] To calculate the standard errors, I cluster on the match-pair level as suggested in Abadie and Spiess (2019); however, I obtain similar levels of precision when clustering at higher levels of aggregation such as the census tract or community area (see Appendix D.2 for more detail).

Since the variation that identifies the event study coefficients emerges from the shifting of gang borders, it is useful to spend some time briefly discussing why gang boundaries move. Broadly speaking, my qualitative work along with the sociological and criminological literature on gangs leads me to believe there are four primary reasons that gang boundaries move. First, gang sets may initially emerge from groups of school-aged children and dissolve after they age out of crime. Thus the age profile of potential gang members may play an important role. Second, gang sets may enter into violent conflict with rival gangs and ultimately be driven out via the resulting attrition. Third, gang sets may be driven out because of police taking actions that explicitly target certain gangs or certain gang members. Fourth, gangs will also expand into areas due to factors related to the neighborhood that make certain blocks

[26]While equation (2) could theoretically suffer from complications related to the conditional variance weighting described earlier, in practice the average effects captured by this specification are what one would expect given a visual inspection of the event study diagrams, so it does not appear to matter in this particular application.

or street corners appealing for the sale of illicit narcotics. I would describe the rest of the variation as idiosyncratic.

This discussion suggests multiple important threats to identification. For example, violent conflict could be the proximate cause of a boundary shift, which would lead me to understate the magnitude of the causal effect via a classic Ashenfelter-type dip (Ashenfelter et al., 1985). In addition, demographic shifts could perhaps be simultaneously causing both changes in the gang boundaries and changes in crime. I could also be that increasing demand for narcotics pulls gangs into new territory where crime would have increased anyway.

There are also important concerns related to crime reporting. First and foremost, the CPD generate these maps as a law enforcement tool. If the boundary shifts in turn change the way police patrol a given neighborhood or report individual crimes, that could lead to changes in reported crime that do not reflect changes in underlying criminal activity as attributable to the gangs. It could also be that neighborhood residents change the way they report crime in response to gang occupation out of fear of retaliation. It is also possible that occupation could cause a reallocation of crime into or out of gang territory when what I would like to measure is the net increase or decrease.

Thus in Section 6.1, I explore a variety of plausible confounding factors and alternative mechanisms. There I show that the findings in this paper cannot be explained by preexisting trends in crime, changes in police surveillance, complicated spatial trends in policing, crime displacement, exposure to public housing demolitions, reporting effects, or demographic trends. I also find that gang occupation has no effect on crimes that have no obvious connection to gang activity. Taken together, the balance of the evidence suggests that the coefficients I recover from equations (1) and (2) warrant a causal interpretation.

Finally, I note that I have also explored gang "exit" effects as well as the impact of gang-to-gang transitions; however, it is much more difficult to produce a compelling counterfactual for these types of border shifts. The qualitative work suggests that while gang entry into unoccupied areas is largely driven by "age-profile" effects, drug market concerns, and other idiosyncratic factors, gang exits and gang-to-gang transitions are in-part driven by gang conflict and directed police action which make them less desirable as a source of identifying variation. For that reason, I have restricted the analysis in this paper to gang entry only.

6 The causal effect of gangs on neighborhood outcomes

Figure 8 displays event study coefficients from model (1) for four primary IUCR outcomes of interest: battery, narcotics violations, prostitution, and robbery. The figures suggest that there is little difference in crime trends between the treated and control blocks prior to gang occupation.[27] However, gang occupation is accompanied by a sharp divergence in the crime trajectories of the treatment and control blocks. Narcotics, batteries, and prostitution all sharply increase, while robberies experience a sharp drop.

[Figure 8 about here.]

Table 2 summarizes these pre-post occupation differences by replacing the event study indicators ($d_{b\tau}$) from model (1) with a single treatment indicator that takes a value of one in all periods after occupation as in model (2). Thus the estimates in Table 2 are generated by effectively comparing the average pre-period event study coefficient to the average post-period event study coefficient. In addition to the four crimes examined in Figure 8, Table 2 also contains estimates for other crimes that gangs might be expected to influence: assault, trespassing, homicide, and weapons violations. See Appendix B.3 for the associated event study diagrams for these additional crimes and Appendix B.4 for the results of formal statistical tests for pre-existing trends. Importantly, I find no evidence that the matched comparison blocks were trending differently than the blocks that experienced a gang occupation prior to the occupation date.

[Table 2 about here.]

As suggested by the event study diagrams, Table 2 shows that battery, narcotics, prostitution, and robbery are all meaningfully impacted by gang occupation. To interpret the magnitude, it is helpful to scale each by the mean of the comparison group. Thus in percentage point terms, the largest effects occur for prostitution (51.9%) and narcotics (18.5%). By way of contrast, the impact on violence is relatively small (6%). The table also shows that gangs cause a moderately sized increase in weapons violations (9.8%) and criminal trespassing (19.6%). The fact that the results are largest for narcotics, prostitution, and trespassing fits well with the qualitative literature about the nature of the open air drug markets where these gangs typically operate. Also, the results on batteries and weapons violations fits well with the summary statistics in Table 1 concerning violence and guns. Thus my preferred interpretation of these results is that they effectively measure the impact of a specific gang-set operating an illegal market.

[27] Appendix B.4 contains the results of formal tests for pre-trends applied to every IUCR crime category contained in the data. I find no evidence that the matched comparison blocks were trending differently than the blocks that experienced gang occupation prior to the occupation date.

The impact on homicide is a fairly precise zero. This is surprising given that we know from both the qualitative work and news reporting that gang-related homicide is real. However, census blocks are reasonably small, and it is not clear that we would expect gang related homicides to occur in the precise blocks where gangs sell illegal narcotics. One of the gang members I interviewed told me that the easiest way to find a rival gang member who they planned to murder was to wait in a car outside of the house belonging to the mother of the rival gang member's children. Thus it may be that homicide is particularly prone to bias emerging from spatial displacement. Consistent with this idea, in Section 6.1, I show that once I aggregate the data to the census tract level, there is some evidence that gangs have a causal effect on homicide.[28]

To obtain a more precise understanding of which crimes are actually responding to gang occupation, I examine the description codes contained in the geospatial crime data. Figure 9 displays regression coefficients from model (2) except that the outcome variable is now a standardized count of crimes that were labeled with the indicated description code. The purpose of standardizing the dependent variable to have a mean of zero and standard deviation of one here is to ensure that the estimated coefficients are comparable across description categories with very different base rates.

Figure 9 shows a profile of causal effects that fits closely with the understanding of gang crime that emerges in the qualitative literature. The increase in battery is disproportionately driven by incidents where a gun or knife is listed as an aggravating factor. The disproportionate increase in domestic violence is perhaps more puzzling; while I cannot say definitively, I would conjecture that it is related to the rise in prostitution. The increase in narcotics is disproportionately driven by the possession and sale of crack cocaine as well as the solicitation and attempted possession of narcotics. The increase in weapons violations is disproportionately driven by incidents involving possession, sale, or use of a handgun, which fits the profile of gang violence detailed in Kapustin et al. (2017). The increase in prostitution is largely driven by increases in public prostitution, and the reductions in robbery come from a decline in strong-arm incidents (i.e., muggings) and robberies involving a gun.

[Figure 9 about here.]

Taken together, the results in Figure 9 and Table 2 suggest that gangs cause small increases in violence, often involving guns, as a result of operating illegal drug markets. Of particular

[28] In general, I prefer the micro-estimates to the more aggregated versions. Leveraging the subcensus tract variation offers three distinct advantages: (1) the occupation "event" is easier to define; (2) it will allow me to explore more fine-grained sources of variation to tease out the role of market structure in Section 7; and (3) the additional variation appears to buy me more precision even after clustering the block-level results at the census tract level.

interest to this interpretation is the decline in robberies, which is consistent with the idea that gang members actually protect their customers. Unfortunately, I did not ask about this possibility during any of my field interviews since I did not have these results yet. However, the drug markets I visited felt safer to me than many other parts of the neighborhoods that we spent time in, as it was very clear that the gang exerted a substantial degree of control over the immediate vicinity of the drug market.

More generally, the results in Figure 9 conform closely to what we might expect to find given the the descriptive statistics presented in Table 1 and given the qualitative accounts in criminology and sociology. In conjunction with Section 6.1, where I probe the most likely threats to identification and conduct a variety of robustness checks, these patterns strongly suggest that a causal interpretation is warranted here.

6.1 Addressing threats to identification and robustness

In this section, I present a variety of robustness checks meant to probe the most plausible threats to identification. Table 3 presents results from model (2) except that I have included different sets of controls. Column (1) contains the main results from Table 2 as a baseline. I now walk through each of the columns and the source of selection it addresses in turn.

The results are not driven by pre-existing trends

One threat to identification is the possibility that the results are driven by pre-existing trends in crime. If the blocks that became occupied were on a different crime trajectory than the blocks that did not, or some sudden change in the block trajectory subsequently caused gang occupation, that could drive a spurious finding. To address this possibility, I will first note that I have used the event study coefficients to test for pre-trends for every IUCR crime category contained in the geospatial data. I find no evidence of differential trends in the pre-period (see Appendix B.4). To further probe this threat to identification, I estimate models that include the interaction of a linear trend with baseline average crime in each of the IUCR crime categories. Column (2) of Table 3 shows the results, which do not differ substantially from the simpler model. Thus pre-existing trends or differential trends driven by baseline differences in crime cannot account for the results.

Demographic change does not appear to be an important source of confounding

Next I explore the possibility that demographic change or gentrification is an important source of confounding. In column (3), I present results from a model that includes controls for the interaction of year 2000 census population counts for each race with a linear trend. The fact that the results are robust suggests the findings are not driven by long-run demographic change. In addition, in the next section, I provide evidence that population at the census tract level is largely unaffected by gang expansion. Together, these finding suggests that demographic change or gentrification are unlikely to be an important source of confounding.

Endogenous changes in police monitoring do not explain the findings

Next I explore the possibility that the findings are caused by changes in police behavior that emerge in response to the boundary shift. Thus in column (4) I include as a control the number of individuals the local police beat adds to the CPD gang database each year. It is possible police activity is causally affected by gang expansion. This fact makes the point estimates in column (4) a bit difficult to interpret, since I am effectively conditioning on an outcome of the treatment. With that limitation in mind, the fact that the results in column (4) are stable is still reassuring. If the effects I document in this section were entirely the product of changes in police behavior driven by gang occupation, I would expect the inclusion of this variable to generate some attenuation in the point estimate. This does not appear to be the case, suggesting endogenous changes in police monitoring are not the primary cause of the findings.

Results are robust to controls for public housing demolitions

Next I consider the possibility that public housing demolitions that occurred over this time period as part of a large urban renewal program could be driving the results. If public housing demolitions caused sudden neighborhood changes that in turn led gangs to enter them, this could result in a spurious finding. To check for this possibility, I use the data from Bruhn (2019) to construct variables that measure the cumulative number of public housing units demolished each year within 0.25 miles, 0.5 miles, and 1 mile of each census block's centroid. Column (5) shows the point estimates that emerge after controlling for these variables. That the results are stable suggest that public housing demolitions are not an important source of confounding.

Results are not driven by different trends in police work

Next, I consider the possibility that there were changes over time in police tactics and procedures that differentially affect the blocks that become occupied. To address this, I measure the distance of each block in the data to every police station in the city[29] and include the interaction of these distances with a linear trend as controls. Column (6) displays the results, which continue to be stable. Thus the results do not appear to be driven by complicated geospatial trends in the nature of police work.

[Table 3 about here.]

Issues related to crime reporting are unlikely to account for the results

Next I consider the possibility that the results are driven by changes in the way neighborhood residents report crime. First, if reporting issues were solely responsible for the pattern of findings in Table 2, it would have to be that gang occupation causes neighborhood residents to become *more* likely to report crimes related to drugs, violence, etc. and *less* likely to report robberies. It is difficult to think of a theory based on reporting effects that would generate such a pattern. In addition, I show that when aggregated to the census tract level, gangs do appear to cause murders, which are less likely to suffer reporting issues.

However, if reporting uniformly decreases across all gang-related crimes as a consequence of occupation, this might lead to understated magnitudes for violence and narcotics while generating a spurious negative finding for robberies. To probe this possibility, I ask whether gang occupation has a causal effect on the tendency for neighborhood residents to call the city of Chicago's Department of Public Works to request graffiti removal.[30] If the tendency for neighborhood residents to call to report a crime is highly correlated with their tendency to call to report graffiti, and gangs do not themselves cause graffiti, then a finding that gang occupation negatively affects the tendency for neighborhood residents to report graffiti would be concerning. Estimates from model (2) suggest that the effect of gangs on graffiti reporting is 0.117 with a standard error of 0.254; however, it is likely that gangs themselves may create graffiti, meaning the test is not dispositive of reporting issues. With that said, the lack of a negative point estimate is still reassuring. Thus taken together, the evidence suggests it is unlikely that the results are entirely driven by reporting effects.

Results are not explained by crime displacement

Next I consider the possibility that the results are driven by crime displacement. For example, it could be that gang occupation does not change the overall amount of crime but instead just

[29] I obtained the grid coordinates of the police stations from publicly available files on the Chicago open data portal.

[30] See Appendix A.4 for a complete discussion of the graffiti removal data.

causes criminals to move activity from one location to another. To address this possibility, I aggregate the sample to the census tract level and estimate model (2).[31] Table 4 contains the results. In all cases, similar results hold across the main outcome variables when considering census tracts, suggesting the results are unlikely to be a pure relocation effect. Further, the table also shows some evidence that gang occupation events cause murders as well.

[Table 4 about here.]

Gang occupation does not cause crimes with no obvious connection to gang activity

As a final sanity check on identification, I ask whether gang occupation causes crimes that the qualitative work suggests are unlikely to be causally affected by street gangs. In Figure 10, I plot standardized regression coefficients similar to Table 9, except that I only consider description codes from crimes that fall under the broad IUCR categories of deception, arson, and kidnapping. The figure suggests that gangs do not cause increases in crimes like electronic harassment, labor theft, or fraud.

[Figure 10 about here.]

Taken together, the evidence in this section suggests that the main results warrant a causal interpretation. While it is never possible to rule out all sources of confounding, it is very reassuring that 1) the specific pattern of findings across crime types corresponds so closely to priors about the ways that gangs might affect their neighborhoods and 2) the results are robust to so many alternative specifications, threats to identification, and robustness checks.

6.2 Gangs negatively impact housing values and median income

I conclude this section by exploring the causal effect of gangs on other neighborhood-level outcomes. Unfortunately, the outcome variables of interest are only available in the ACS at the census tract level and in the form of five-year averages, making it complicated to directly leverage the variation I exploited at the census block level. Thus in the interest of lining up the analysis in this section as closely as I can with the variation used in the preceding sections, I use a matched and stacked 2x2 difference-in-difference design for these outcomes, where the matches are defined using the block-level matches I described earlier.

[31] The aggregation procedure I use in this section is identical to the procedure used in 6.2 except that, unlike the ACS outcomes, I observe crime for many years before and after the gang occupation event. For that reason, I can use the aggregated data to directly estimate model (2) rather than the more parsimonious stacked 2x2 dif-in-dif specification I apply to the ACS data in Section 6.2.

To see how this works, it is easiest to start with a concrete example. Consider a matched pair of census blocks where the occupation event occurs in 2012. For these blocks, I take the five-year ACS average for the corresponding census tracts from 2007 to 2011 and call that the pre-period. I also take the five-year ACS average for the corresponding census tracts from 2012 to 2017 and call that the post-period. I then repeat this procedure for every census block match pair in the data. This procedure defines a one-to-many match between treated census tracts (i.e., those that experience an occupation event) and the census tracts that contain the control blocks. In many cases, the resulting matches turn out to be one-to-one. When they are not, I reduce them to a one-to-one match by randomly choosing which of the implied control census tracts to use as the matched comparison.[32] From there, I estimate models of the following form:

$$y_{ct} = \alpha + \omega d_c + \pi p_t + \beta d_c \times p_t + \eta_{ct}, \tag{3}$$

where y_{ct} is the ACS outcome for census tract c during the pre-period ($t = 0$) or the post-period ($t = 1$), d_c is an indicator for whether census tract c ever experiences an entry event, and p_t is an indicator that takes a value of 1 during the post-period ($t = 1$). I also estimate models that directly mimic specification (2) by replacing the first two terms (d_c and p_t) with census tract and match pair by post-period fixed effects. Table 5 contains the results.

[Table 5 about here.]

From Table 5, there appears to be moderate effects on household income (3.1% relative to the control mean) and median home values (3.2% relative to the control group mean). Other outcomes, such as per capita income and total population, do not appear to be significantly affected. However, I caution that the evidence in Table 5 should be taken as suggestive (not definitive). All of these estimates are much less precise than those found for crime. Further, the nature of the data makes it impossible to directly explore pre-existing trends for these outcomes, since there are no pre-period data. That said, the direction of the impacts are consistent with what we would expect given the findings for crime.

[32] Reducing the data to a one-to-one match is attractive because it ensures the estimated treatment effect will correspond to a census tract average rather than to a weighted average based in part on the number of control observations within each match group.

7 Gang impact and the organization of the illegal market

In this section, I explore the connection between the production of crime and the industrial organization of the black market. Theoretically, it can be socially optimal to grant market power to a criminal organization under two sets of conditions. First, it could be that having a single gang or a small number of large gangs leads to anti-competitive behavior. In the case of a monopoly or a collusive oligopoly, this can lead to restrictions on the quantity of crime supplied and hence a reduction in the externality attributable to the social "bad." On the other hand, even if the crime in question poses no externality, market power in the hands of a criminal organization can still be beneficial if the supply of law enforcement is endogenous. This is because a gang with market power is better positioned to internalize costly police effort (Buchanan, 1973).

To see this formally, consider a simple model of Cournot competition in the black market. Suppose there are G identical gangs indexed by g that each produce a quantity of an illegal narcotic q_g. Let $Q = \sum_G q_q$ denote market supply. For simplicity, assume that demand is linear so that in equilibrium, price will be a linear function of Q given by $P = a - bQ$. Assume further that the marginal cost of production to each gang is a linear function of the total supply of law enforcement: $MC = c + dL$, where $L = \varepsilon \frac{Q}{G}$, with ε representing the elasticity of law enforcement effort to the supply of the illegal narcotic.

Under these conditions, it is simple to show that the equilibrium supply of the illegal narcotic will be

$$Q^* = \left(\frac{G}{G+1}\right)\left(\frac{a-c}{b+\frac{d\varepsilon}{G}}\right) \tag{4}$$

This model makes two sets of predictions I explore in the data. First, observe that the equilibrium supply of the illegal narcotic is increasing in the number of gangs. Further, it is also the case that when the market is competitive,[33] the marginal effect of an additional gang will be small. Thus if we observe that gangs cause larger causal effects in more competitive environments, we can reject simple market-based explanations for criminal behavior. Second, observe that when $d > 0$ so that gangs internalize law enforcement effort, Q^* is declining in the elasticity of law enforcement (ε). Thus we expect that the causal effect of adding an additional gang to the market will be larger when the elasticity of law enforcement effort is smaller. If we find that the causal effect of adding an additional gang to the market does not

[33]Formally, as $G \to \infty$, $Q^* \to \frac{a-c}{b}$ so that equilibrium quantity no longer depends on the number of gangs.

vary with the elasticity of law enforcement, this would again suggest that we need a more complicated model to explain the data.

To test the first prediction, I estimate models using the same sample and fixed effects as described for model (2), but now I also allow the causal effect of gang occupation to vary with other observables related to gang competition. Table 6 shows the results. Columns (1) and (5) present the baseline estimates from Table 2 as a reference, while columns (2) and (6) interact the post-gang indicator with the number of gang sets and members found in the 1982 birth cohort.[34] Thus I wish to determine whether gangs of similar size as measured by membership have larger causal effects when they experience more competition due to having more independent groups (also known as gang sets) operating within their territory. While small, the point estimates suggest that gangs that experience more internal competition as a result of gang fracturing generate more narcotics violations and more batteries.

In columns (3) and (7) I replace the post-event indicator with an alternative set of indicators that correspond to the number of gangs within a quarter mile of the census block. The "one gang" variable takes a value of 1 when there is exactly one gang present in the block. The "two+ gangs" variable takes a value of one when there are two or more gangs present in the block. The point estimates suggest that supply is increasing in the number of gangs but the marginal effect of adding the second gang is larger than the first. Finally, in columns (4) and (8), I replace the post-event indicator with indicators for whether the centroid of the block is close to gang territory, on the border of gang territory, or deep inside gang territory.[35] If blocks on the border experience more competition, we would expect to see smaller increases in crime there than we do deeper inside gang territory. Instead, we find that the effects on the border are larger.

[Table 6 about here.]

In particular, across all three specifications I find that the variation in competition is particularly important for narcotics violations. This is what we would expect to see if the end point of sale and distribution of illegal narcotics is the primary gang industry. However, we also see similar heterogeneity in batteries. There are two potential explanations for this finding. First, violence could be a cost gangs incur when they compete with other gangs similar

[34]I choose this birth cohort because I wish to measure the degree of fracturing "at-baseline." Since gang members are young (see Table 1), I decided to pick 2002 (which is shortly before the first year where I observe the maps) and subtract 20 years. However, the results are not sensitive to this choice.

[35]More precisely, let d denote the distance of the block centroid from the nearest gang border, with negative values denoting a centroid that is inside gang territory and positive values denoting a centroid that is outside gang territory. I call a block near if $.125 < d < 0.25$, on the border if $-0.125 < d < 0.125$ and deep if $d < -.125$. I chose these thresholds based on the visual evidence in Figure 4 regarding where crime shifts relative to the border.

to how firms in legal markets compete with one another via costly advertising campaigns. Alternatively, it could be that gangs operating near one another generate more opportunity for conflict to emerge from social interactions. This second explanation is consistent with the qualitative literature (Aspholm, 2020), the literature on behavioral interventions to reduce violence (Heller et al., 2017), my own experiences interviewing gang members, and the summary statistics from Section 4 that show race as an important predictor of gang territory and conflict. Thus my preferred interpretation of the results concerning batteries in Table 6 is that competition produces environments that facilitate social interactions that in turn lead to violence. However, an important limitation of this work is that I cannot definitely distinguish between these two possibilities.

To explore the role of the endogenous police response, I estimate how gang causal effects vary with distance to the nearest police station. If it is more costly for police to monitor drug markets that are further away (so that the elasticity of law enforcement effort is declining with distance), then this interaction term will indirectly measure how the causal effect of gang entry varies with the elasticity of the law enforcement response.

[Table 7 about here.]

Table 7 displays the results. For virtually all crimes, the interaction between gang occupation and distance to police is either close to zero or negative. This finding is inconsistent with the hypothesis that gangs respond to or can internalize police effort. However, it is important to point out that this analysis is limited for two reasons: (1) the interaction terms are imprecisely estimated, and (2) I do not directly observe the elasticity of police effort (only a plausible shifter). However, the fact that virtually all the coefficients on the interaction terms are close to zero is highly suggestive.

8 Conclusion

In this paper, I use novel data on the geospatial distribution of gang territory in Chicago and its evolution over a 15-year period to document new stylized facts regarding the distribution and dynamics of gang territory. I use a matched event study design to estimate the causal effect of gangs on crime and other neighborhood-level outcomes and find evidence that gangs cause small increases in violence in highly local areas due to conflict related to illegal markets. I also find evidence that gangs reduce neighborhood income and property values. Finally, I explore how these causal effects vary with the industrial organization of the illegal markets. The data are not explained well by simple theories of competition in the black market, suggesting behavioral factors and social interactions play an important role in the production of gang violence.

These findings are important for two reasons. First, estimating the causal effect of gangs is a necessary ingredient for any rational policy response that seeks to weigh the costs and benefits of various gang interventions. Second, the results help to shed light on the limitations of standard economic theories of human behavior as a tool for understanding illegal markets. While there are likely many settings where such an approach will prove useful, Chicago does not appear to be one of them.

Understanding the types of crimes, settings, and institutions where standard economic theories do and do not apply to criminal markets is an important area for future work. I would conjecture that the age and early life experiences of the criminals involved plays an important role. For example, Blattman and Annan (2010) document the negative impact of soldiering on the mental health and subsequent life outcomes of children. More recently, Jácome (2021) finds that losing access to Medicaid, and in particular behavioral health services, drives children into criminal behavior. Perhaps exposure to violence, drugs, and poverty at a young and malleable age leads one to become a criminal whose behavior cannot be accurately characterized as rational. This is consistent with my data, where the gang members are young and age into crime against the backdrop of a harsh, violent, and poverty-stricken environment.

An important limitation of this study is its inability to provide compelling evidence regarding the long-run impact of organized crime. It is possible that criminal organizations play an important role in the intergenerational reproduction of poverty and violence via their influence on neighborhood institutions and culture. Indeed, many of the same neighborhoods occupied by gangs in my data are the ones that were explored by Frederick Thrasher in his seminal work on Chicago youth gangs in the 1920s (Thrasher, 2013). In that sense, the estimates in this paper are inherently short run in nature and may understate the true costs that gangs impose. However, developing this richer, long-run view of the causal effect of gangs is left for future work. More broadly, better understanding the specific neighborhood-level factors that produce such environments should be a critical goal for subsequent empirical work.

References

Abadie, A. and Spiess, J. (2019). Robust Post-Matching Inference. Technical report. 18, 47, 50

Acemoglu, D., De Feo, G., and De Luca, G. (2017). Weak States: Causes and Consequences of the Sicilian Mafia. 6

Arora, A. (2018). Juvenile Crime and Anticipated Punishment. *SSRN Electronic Journal*. 3

Ashenfelter, O., Card, D. E., Ashenfelter, O., and Card, D. (1985). Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs. *The Review of Economics and Statistics*, 67(4):648–60. 19

Aspholm, R. (2020). *Views from the Streets: The Transformation of Gangs and Violence on Chicago's South Side*. Columbia University Press, United Kingdom. 2, 4, 5, 10, 12, 13, 16, 29

Barone, G. and Narciso, G. (2015). Organized Crime and Business Subsidies: Where Does the Money Go? *Journal of Urban Economics*, 86:98–110. 6

Bayer, P., Hjalmarsson, R., and Pozen, D. (2009). Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections. *Quarterly Journal of Economics*, 124(1):105–147. 3

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2):169–217. 3

Billings, S. B., Deming, D. J., and Ross, S. L. (2019). Partners in Crime. *American Economic Journal: Applied Economics*, 11(1):126–150. 3

Blattman, C. and Annan, J. (2010). The Consequences of Child Soldiering. *Review of Economics and Statistics*, 92(4):882–898. 30

Blattman, C., Duncan, G., Lessing, B., and Tobon, S. (2019). Gangs of Medellin: How Organized Crime is Organized. *Unpublished Manuscript*. 5

Blattman, C., Duncan, G., Lessing, B., and Tobon, S. (2021). Gang rule: Understanding and Countering Criminal Governance. 5

Brantingham, J. P., Tita, G. E., Short, M. B., and Reid, S. E. (2012). The Ecology of Gang Territorial Boundaries. *Criminology*, 50(3):851–885. 6

Brown, Z. Y., Montero, E., Schmidt-Padilla, C., and Sviatschi, M. M. (2020). Market Structure and Extortion:Evidence from 50,000 Extortion Payments. *NBER Working Paper 28299*. 3

Bruhn, J. (2018). Crime and Public Housing. *SSRN Working Paper*. 5

Bruhn, J. (2019). *The Consequences of Sorting for Understanding School Quality*. PhD thesis. 13, 23

Buchanan, J. M. (1973). A Defense of Organized Crime. In *The Economics of Crime and Punishment*, pages 119–132. 3, 6, 27

Calamunci, F. M. (2021). What Happens in Criminal Firms after Godfather Management Removal? Judicial Administration and Firms' Performance. *Small Business Economics*. 6

Callaway, B. and Sant'Anna, P. H. C. (2018). Difference-in-Differences with Multiple Time Periods and an Application on the Minimum Wage and Employment. *SSRN Electronic Journal*. 18

Carvalho, L. S. and Soares, R. R. (2016). Living on the Edge: Youth Entry, Career and Exit in Drug-Selling Gangs. *Journal of Economic Behavior & Organization*, 121:77–98. 5

Chicago Crime Commission (2019). Chicago Crime Commission. 9

City of Chicago Office of the Inspector General (2019). Review of the Chicago Police Department's "Gang Database". Technical report, City of Chicago, Chicago. 1, 7

Cook, P. J., Ludwig, J., Venkatesh, S., and Braga, A. A. (2007). Underground Gun Markets. *The Economic Journal*, 117(524):588–618. 5

Daniele, G. and Dipoppa, G. (2017). Mafia, Elections and Violence against Politicians. *Journal of Public Economics*, 154:10–33. 6

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9):2964–2996. 18

De Feo, G. and De Luca, G. D. (2017). Mafia in the Ballot Box. *American Economic Journal: Economic Policy*, 9(3):134–167. 6

De Martiis, A. (2020). The Effects of Organized Crime on Distressed Firms. 6

Dell, M. (2015). Trafficking Networks and the Mexican Drug War. *American Economic Review*, 105(6):1738–1779. 5

Dooley, B. D., Seals, A., and Skarbek, D. (2014). The Effect of Prison Gang Membership on Recidivism. *Journal of Criminal Justice*, 42(3):267–275. 5

Durlauf, S. N. and Ioannides, Y. M. (2010). Social Interactions. *Annual Review of Economics*, 2:451–478. 3

Durlauf, S. N. and Tanaka, H. (2008). Understanding Regression Versus Variance Tests for Social Interactions. *Economic Inquiry*, 46(1):25–28. 3

Glaeser, E. L., Sacerdote, B., Scheinkman, J. A., and Scheinkman, J. A. (1996). Crime and Social Interactions. *The Quarterly Journal of Economics*, 111(2):507–548. 3

Goodman-Bacon, A. (2019). Difference-in-Difference with Variation in Treatment Timing. *Unpublished Manuscript*. 18

Great Cities Institute (2019). The Fracturing of Gangs and Violence in Chicago: A Research-Based Reorientation of Violence Prevention and Intervention Policy. In *The Fracturing of Gangs Conference*. 2, 4, 5

Grogger, J. (2002). The Effects of Civil Gang Injunctions on Reported Violent Crime: Evidence from Los Angeles County. *The Journal of Law and Economics*, 45(1):69–90. 5

Hagedorn, J. and Macon, P. (1998). *People and Folks: Gangs, Crime and the Underclass in a Rustbelt City*. Lake View Press, United States. 4

Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., and Pollack, H. A. (2017). Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago. *The Quarterly Journal of Economics*, 132(1):1–54. 3, 29

Herrnstadt, E., Heyes, A., Muehlegger, E., and Saberian, S. (2018). Air Pollution as a Cause of Violent Crime: Evidence from Los Angeles and Chicago. *Unpublished Manuscript*. 11

Jacob, B. A. and Lefgren, L. (2003). Are Idle hands the Devil's Workshop? Incapacitation, Concentration, and Juvenile Crime. *American Economic Review*, 93(5):1560–1577. 3

Jácome, E. (2021). Mental Health and Criminal Involvement: Evidence from Losing Medicaid Eligibility. *Working paper*. 30

Kapustin, M., Ludwig, J., Punkay, M., Smith, K., Speigel, L., and Welgus, D. (2017). Gun Violence in Chicago, 2016. Technical report, University of Chicago Urban Labs, Chicago. 4, 21

Kirby, K. C., Eghigian, M., and Petrenko, J. (2006). The Gang Book. Technical report, Chicago Crime Commission, Chicago. 1, 7, 9, 15, 42, 2

Leeson, P. T. (2007). An-arrgh-chy The Law and Economics of Pirate Organization. *Journal of Political Economy*, 115(6):1049–1094. 5

Levitt, S. D. and Venkatesh, S. A. (2000). An Economic Analysis of a Drug-Selling Gang's Finances. *Quarterly Journal of Economics*, 115(3):755–789. 1, 5

Levitt, S. D. and Venkatesh, S. A. (2001). Growing Up in the Projects: The Economic Lives of a Cohort of Men Who Came of Age in Chicago Public Housing. *American Economic Review*, 91(2):79–84. 1, 5

Lindquist, M. J. and Zenou, Y. (2019). Crime and Networks: Ten Policy Lessons. *Oxford Review of Economic Policy*, 35(4):746–771. 6

Lonsky, J. (2019). Gulags, Crime, and Elite Violence: Origins and Consequences of the Russian Mafia. pages 1–53. 6

Marcolongo, G. (2020). Organized Crime, Earthquakes and Local Public Procurement: Evidence from Italy. 6

Melnikov, N., Schmidt-Padilla, C., and Sviatschi, M. M. (2019). Gangs, Labor Mobility, and Development. *Unpublished Manuscript*. 5

Miller, D. L., Ama Shenhav, N. ., and Grosz, M. Z. (2019). Selection into Identification in Fixed Effects Models, with Application to Head Start. *Working Paper*. 18

Monteiro, J. and Rocha, R. (2017). Drug Battles and School Achievement: Evidence from Rio de Janeiro's Favelas. *The Review of Economics and Statistics*, 99(2):213–228. 5

National Gang Center (2019). National Youth Gang Survey Analysis. 1, 4

Olken, B. A. and Barron, P. (2009). The simple economics of extortion: Evidence from trucking in aceh. *Journal of Political Economy*, 117(3):417–452. 3

Papachristos, A. V. (2009). Murder by Structure: Dominance Relations and the Social Structure of Gang Homicide. *American Journal of Sociology*, 115(1):74–128. 1, 6, 7, 15, 16

Papachristos, A. V., Hureau, D. M., and Braga, A. A. (2013). The Corner and the Crew: The Influence of Geography and Social Networks on Gang Violence. *American Sociological Review*, 78(3):417–447. 1, 6, 15, 16

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1). 1

Pinotti, P. (2015a). The Causes and Consequences of Organised Crime: Preliminary Evidence Across Countries. *The Economic Journal*, 125(586):F158–F174. 6

Pinotti, P. (2015b). The Economic Costs of Organised Crime: Evidence from Southern Italy. *The Economic Journal*, 125(586):F203–F232. 6

Pinotti, P., Decarolis, F., Vannutelli, S., and Fisman, R. (2020). Rules, Discretion, and Corruption in Procurement: Evidence from Italian Government Contracting. *SSRN Electronic Journal*, pages 1–56. 6

Roth, J. and Sant'anna, P. H. C. (2021). When Is Parallel Trends Sensitive to Functional Form? . *Working paper*. 18

Sanburn, J. and Johnson, D. (2017). Chicago's Deadly 2016: See It in 3 Charts. *Time*. 4

Skarbek, D. (2010). Putting the Con into Constitutions: The Economics of Prison Gangs. *Journal of Law, Economics, and Organization*, 26(2):183–211. 5

Sobrino, F. (2019). The Cartel Wars: Fighting Over Opioid Crisis Profits. *Unpublished Manuscript*. 5

Sun, L. and Abraham, S. (2020). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*. 18

Sviatschi, M. M. (2019a). Making a Gang: Exporting US Criminal Capital to El Salvador. *Unpublished Manuscript*. 5

Sviatschi, M. M. (2019b). Making a Narco: Childhood Exposure to Illegal Labor Markets and Criminal Life Paths. *Unpublished Manuscript*. 5

Taniguchi, T. A., Ratcliffe, J. H., and Taylor, R. B. (2011). Gang Set Space, Drug Markets, and Crime around Drug Corners in Camden. *Journal of Research in Crime and Delinquency*, 48(3):327–363. 6

Thrasher, F. M. (2013). *The Gang: A Study of 1,313 Gang in Chicago*. University of Chicago Press. 6, 30

United Nations Office on Drugs and Crime (2005). 2005 World Drug Report. Technical report. 1

Wildeman, C. and Roberto, E. (2015). Tragic, but Not Random: The Social Contagion of Nonfatal Gunshot Injuries. *Social Science & Medicine*, 125:139–150. 6

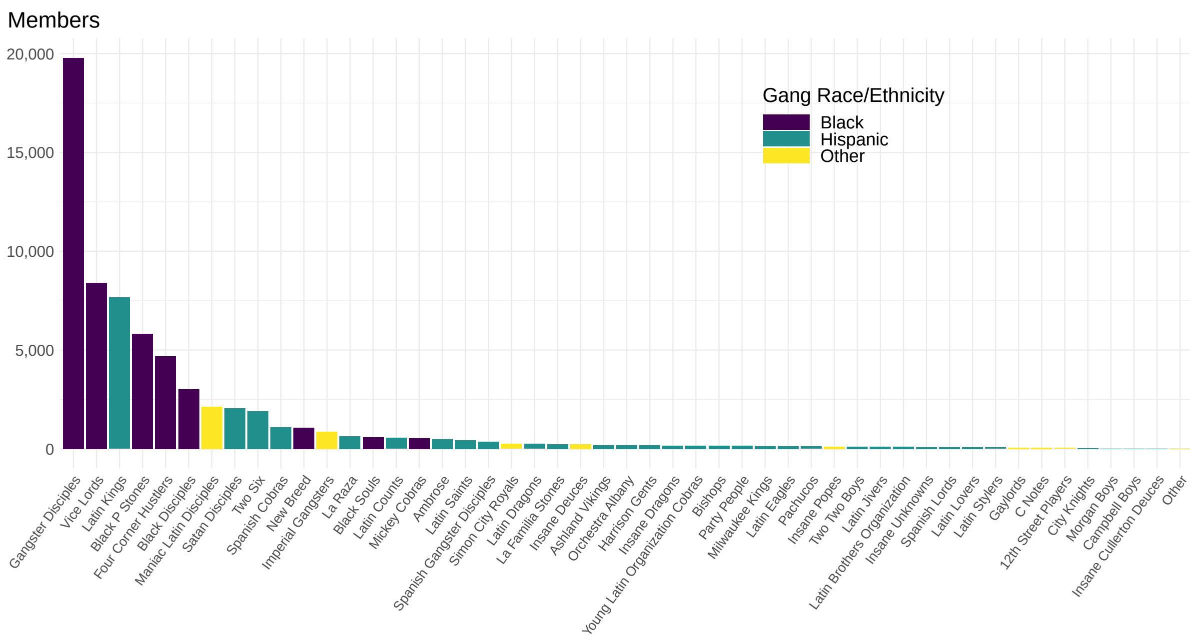

FIGURE 1: Chicago has a large number of small gangs

Note: This figure displays estimated membership for every gang found in the map data in 2004. To proxy for membership, I use the anonymized CPD gang database to count the number of individuals within each gang who belong to the 1975–1990 birth cohorts. The logic behind this proxy is that individuals in those birth cohorts would range in age from 14 to 29 as of the year 2004. However, the diagram is very similar and yields the same broad conclusions if I measure gang "size" using territory area instead of membership. See Appendix B.1

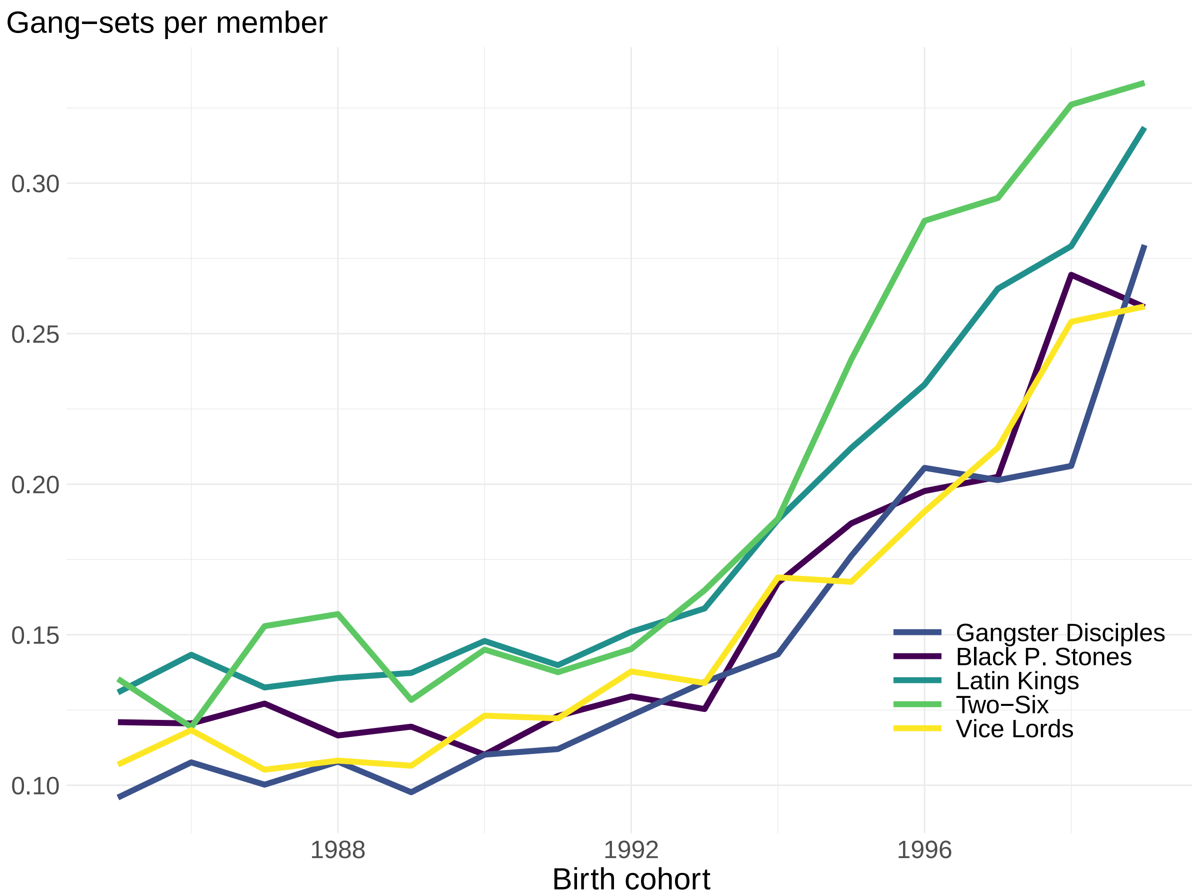

FIGURE 2: Chicago's gangs have fractured over time

Note: This figure plots a measure of gang fracturing for the five largest gangs in Chicago. Fracturing is measured by counting the distinct number of gang sets found in the CPD gang database for each gang and in each birth cohort and then dividing the number of sets by the total number of gang members found in each gang in each birth cohort.

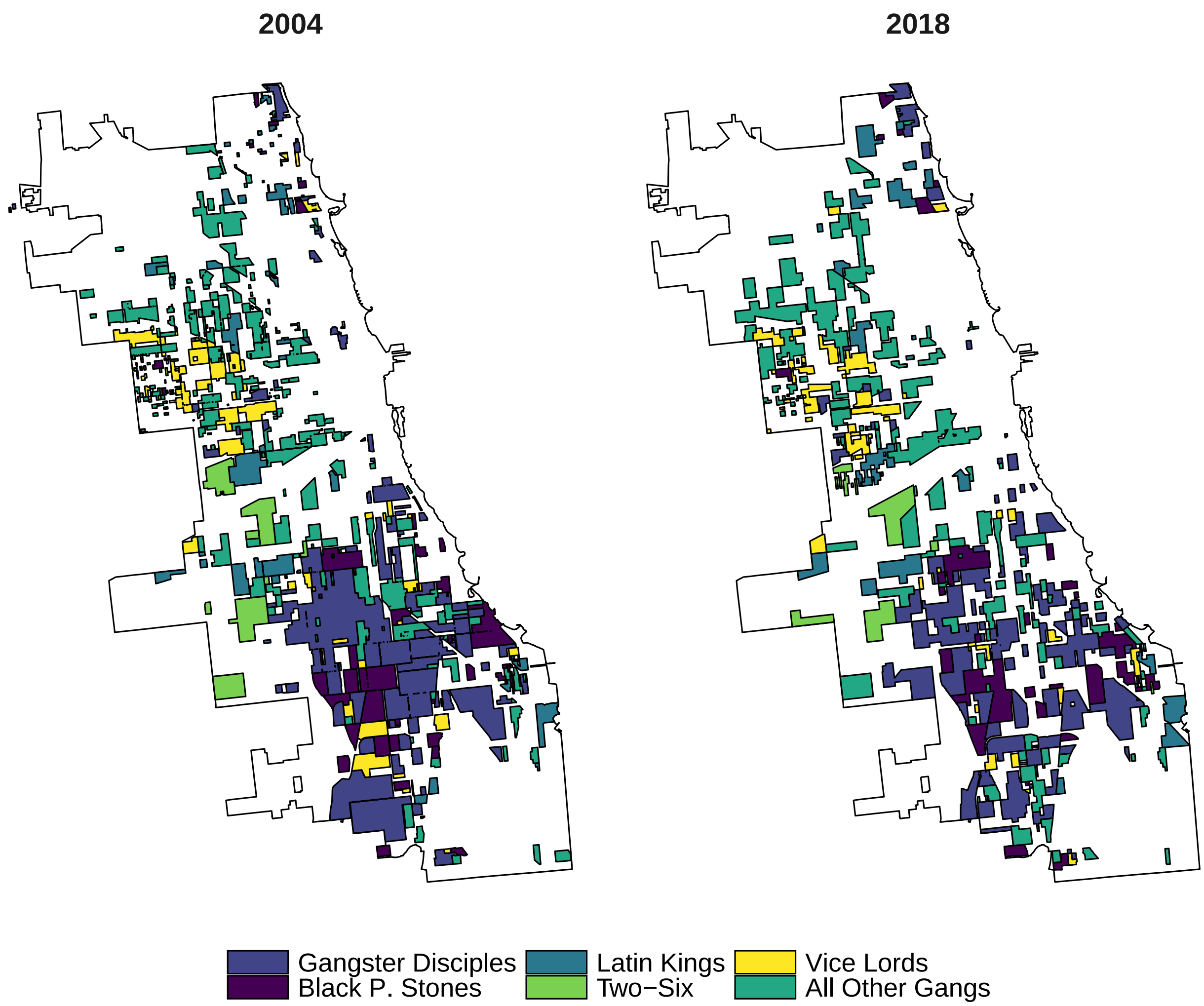

FIGURE 3: Gang Territory in 2004 and 2018

Note: This figure plots maps of gang territory in Chicago from the 2004 and 2018 CPD shapefiles. The gangs with the five largest polygons in 2004 are highlighted. All remaining gang territories are included in the "All Other Gangs" category.

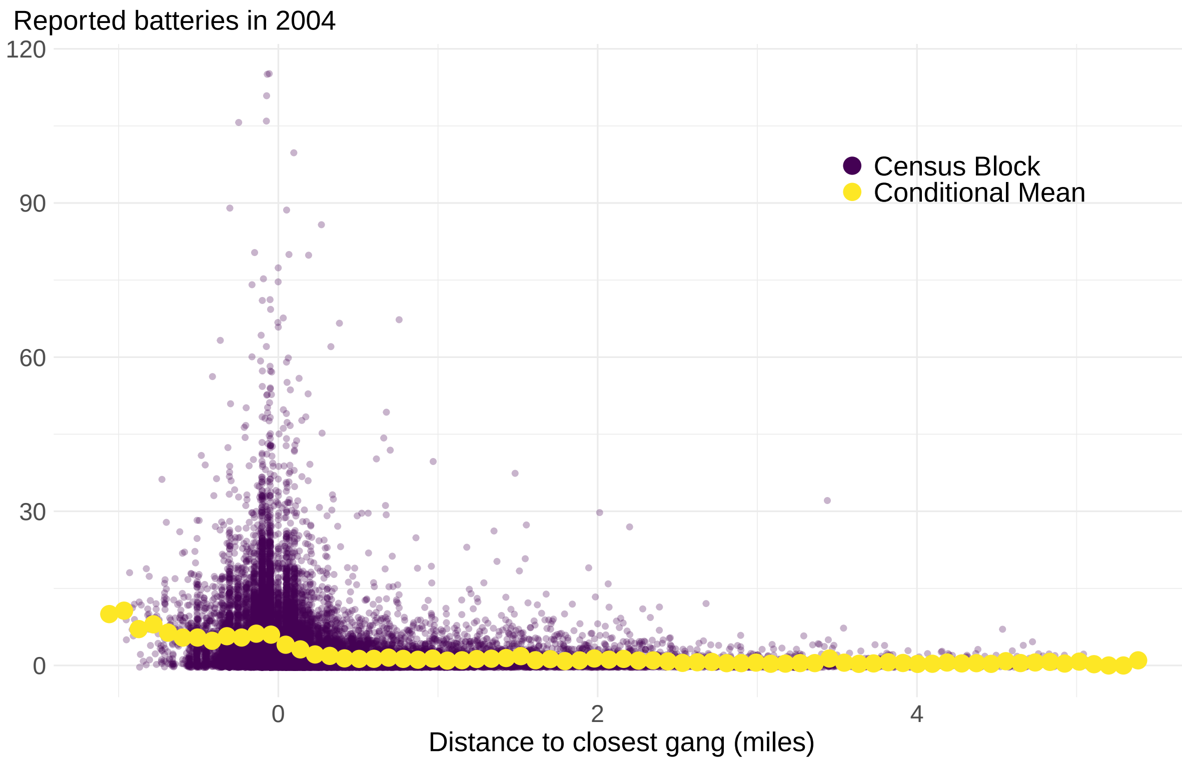

FIGURE 4: Violence increases sharply at gang borders

Note: This figure plots counts of reported batteries against distance to the nearest gang border. Small points with dark shading denote individual census blocks. Large, lightly shaded points denote conditional means within 70 equal-length distance bins. Distance ($\tilde{d}_{bg}$) to each gang (g) for each block (b) is defined as the length of the smallest straight line connecting block b's centroid to gang g's border. Negative values denote centroids that are located inside gang g's territory, and positive values denote centroids that are outside gang g's territory. The distance to the closest gang for block b is then defined as the minimum over all these gang specific distances ($d_b = \min_{g \in G} \tilde{d}_{bg}$).

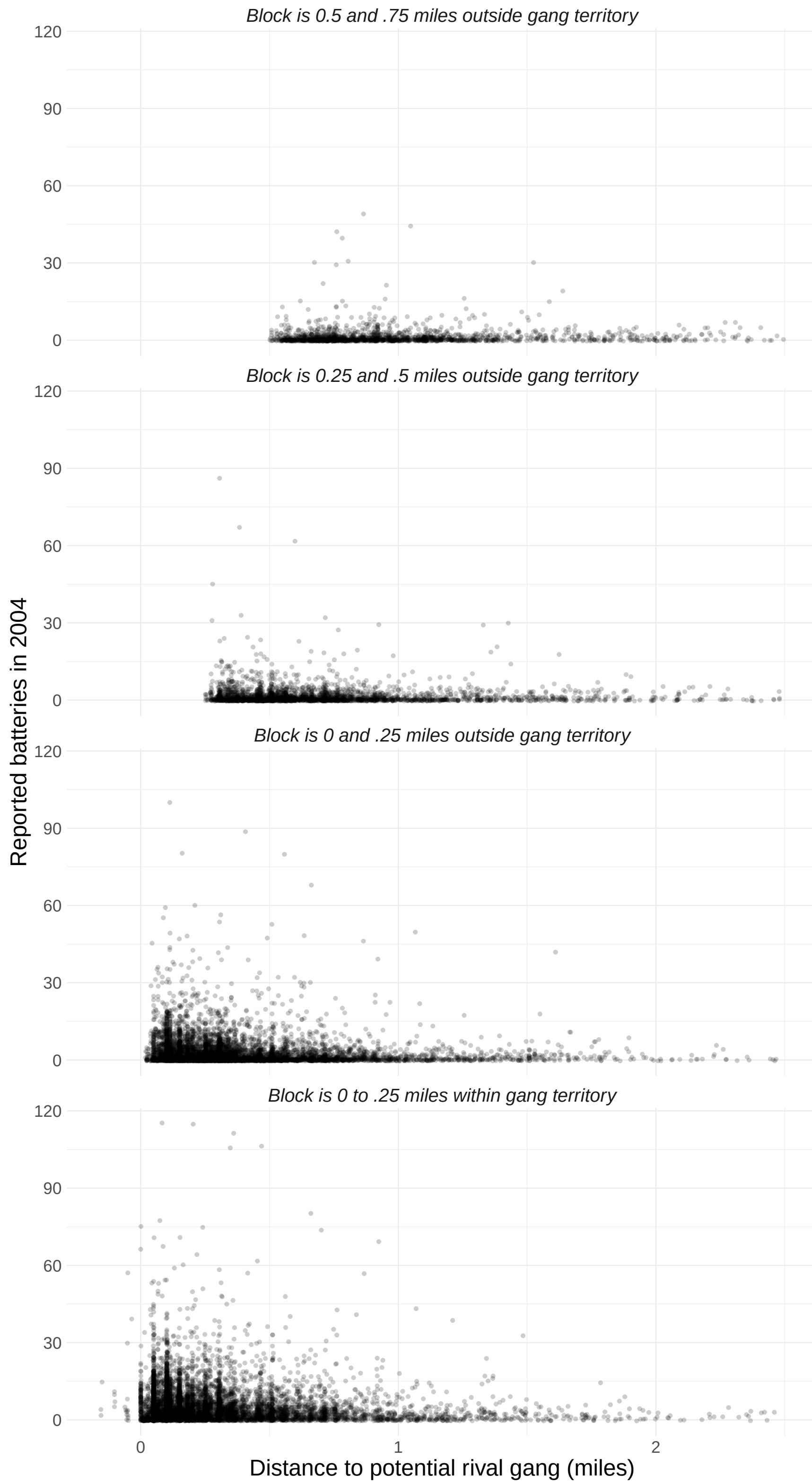

FIGURE 5: Differences in violence emerge only when two gangs are in close proximity

Note: This figure plots reported batteries against distance to the border of a potential rival gang. Each point is a census block observed in 2004. Subpanels condition the sample on distance to the closest gang. Thus all observations in the top panel fall between 0.5 and 0.75 miles outside of gang territory, and all blocks in the bottom panel are 0 to 0.25 miles *inside* gang territory. Distance ($\tilde{d}_{bg}$) to each gang (g) for each block (b) is defined as the length of the smallest straight line connecting block b's centroid to gang g's border. Negative values denote centroids that are located inside gang g's territory, and positive values denote centroids that are outside gang g's territory. The distance to the closest gang for block b is then defined as the minimum over all these gang specific distances ($d_b = \min_{g \in G} \tilde{d}_{bg}$). Distance to a potential rival is the second smallest of these values.

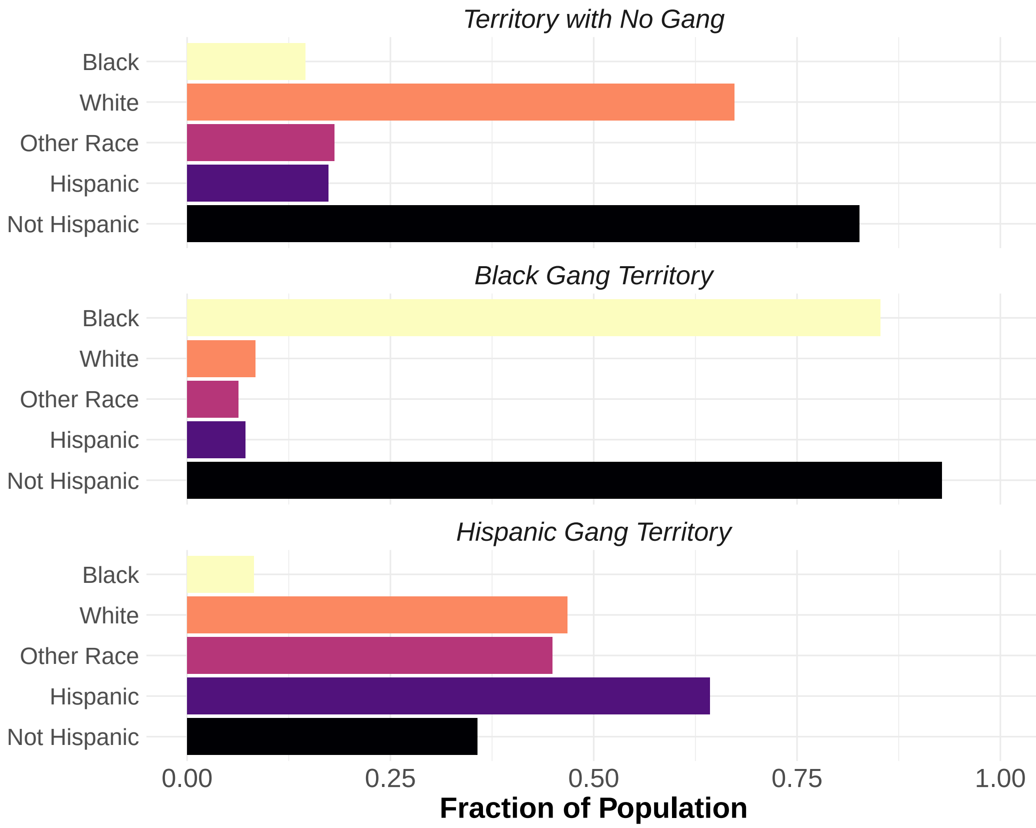

FIGURE 6: Gang race is a strong predictor of territory demographics

Note: This figure describes the racial composition of territory occupied by gangs identified with different races/ethnicities. Gang territory is measured using the polygons found in the 2004 CPD gang map shapefiles. The racial composition of a block is measured according to the 2000 decennial census. The strip text denotes the primary race/ethnicity of the occupying gang as indicated in the 2006 edition of the Chicago Crime Commissions "Gang Book." Bars are shaded according to the race/ethnicity categories denoted on the y-axis to facilitate visual comparisons across subfigures.

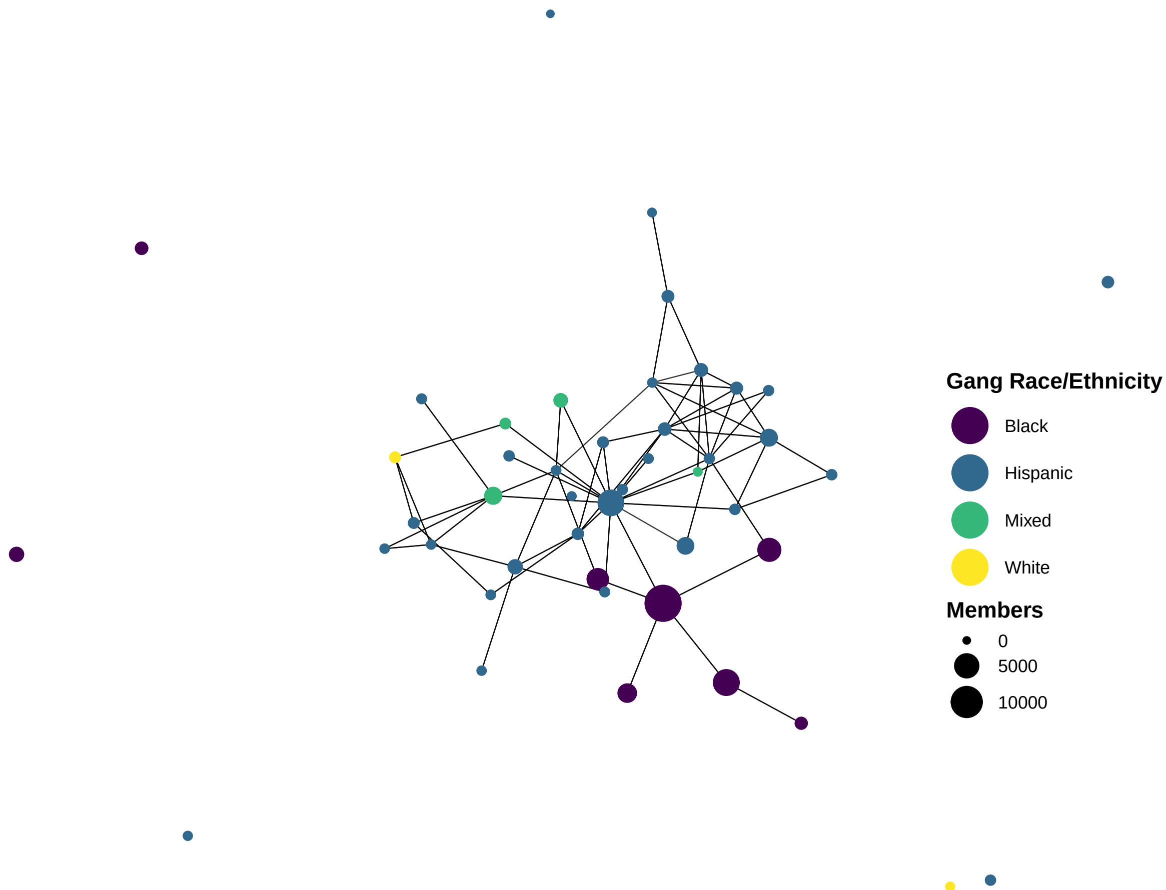

FIGURE 7: Historically, conflict tended to occur between gangs of similar race

Note: This figure describes the historical distribution of rivalries and wars among gangs. Each vertex represents a gang. The size of the vertex corresponds to gang membership as measured via the CPD gang database over the 1975–1990 birth cohorts. The shading of the vertex denotes the primary race/ethnicity associated with the gang as described in Kirby et al. (2006). Edges signify that at one point in time, the gangs were actively at war or considered each other rivals as described in the narrative histories contained in Kirby et al. (2006). Gangs that did not appear in Kirby et al. (2006) are not included in this figure.

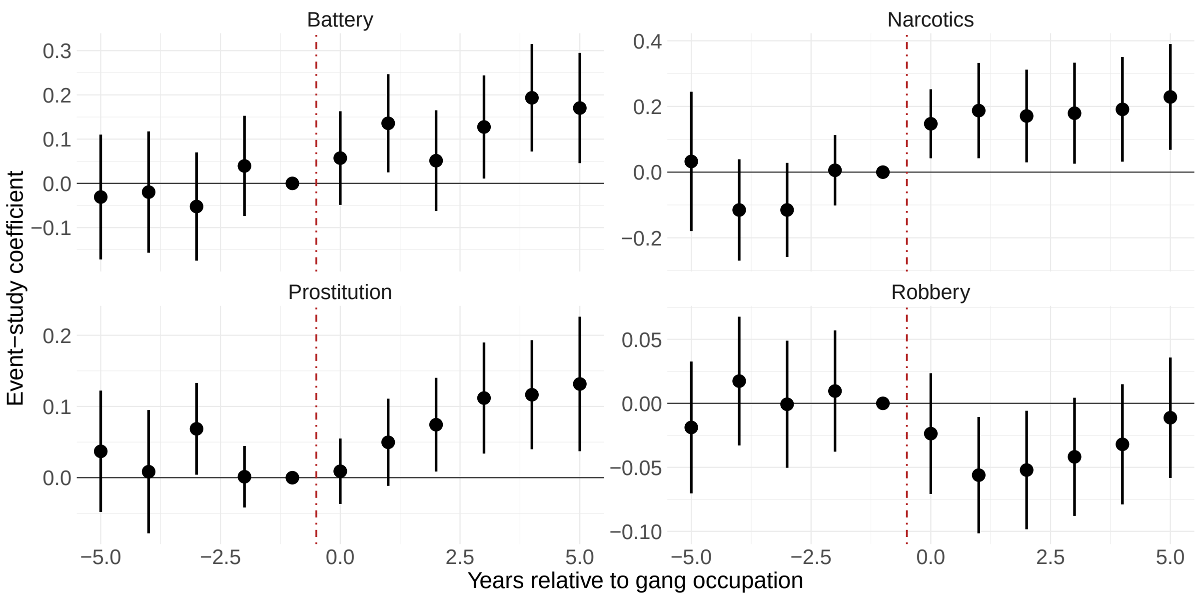

Figure 8: Gang occupation predicts sharp changes in crime

Note: This figure displays event study coefficient from regression (1) for four primary IUCR outcomes of interest: battery, narcotics violations, prostitution, and robbery.

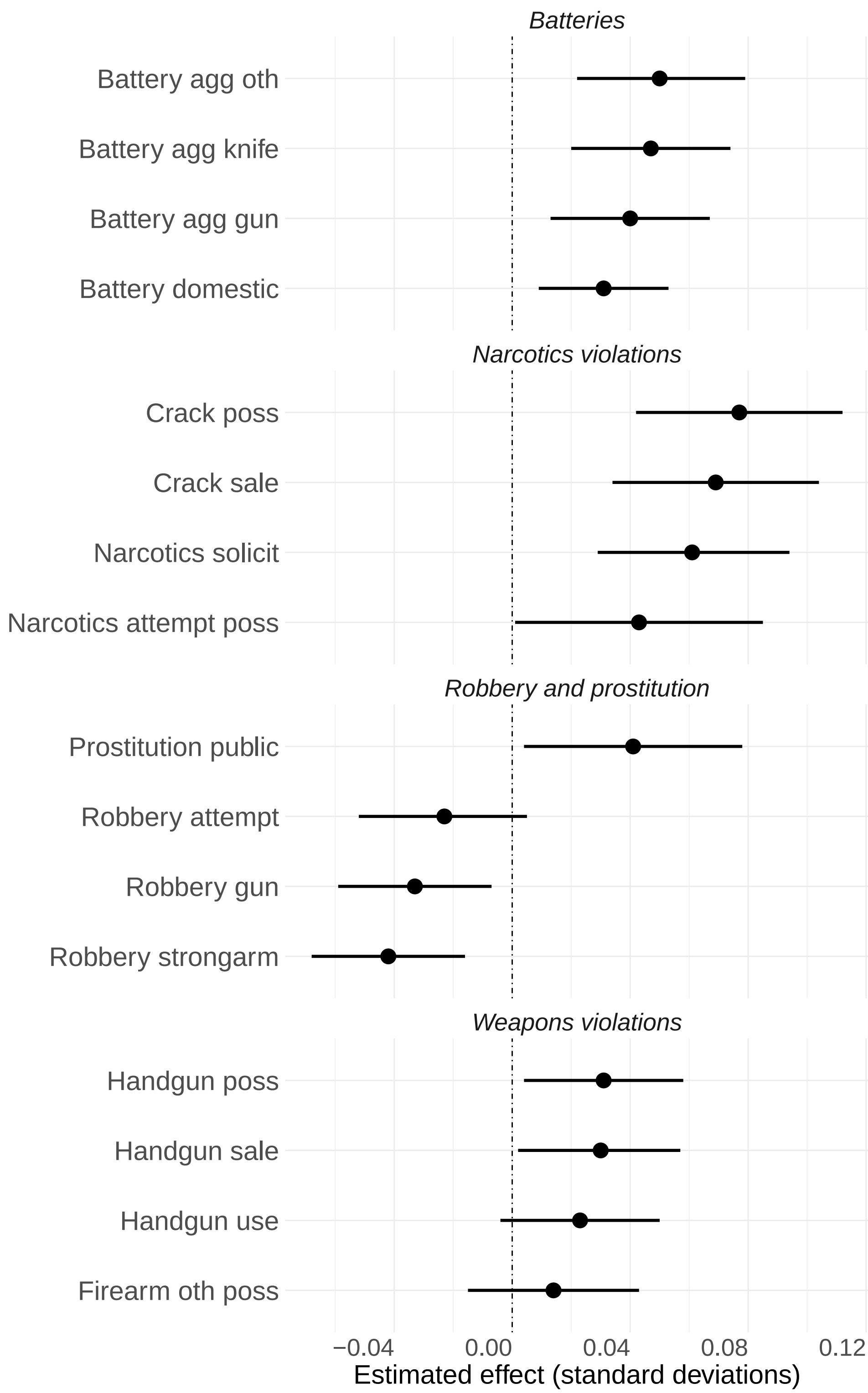

Figure 9: Average effect of gangs on select crime descriptions

Note: This figure presents standardized regression coefficients from model (2) for crimes that had the description label indicated on the y-axis. The strip text denotes the corresponding IUCR code that the crime fell under. The bars represent two standard error intervals. There are over 500 distinct description codes in the raw data, which I aggregate into ≈ 138 categories by fixing spelling errors and lumping very similar crimes together. For visual clarity, I only present the four that had the largest standardized regression coefficient within each IUCR category. The exception to this inclusion criteria is prostitution (which only has two description labels) and robbery (which only has five), which I have combined into a single panel.

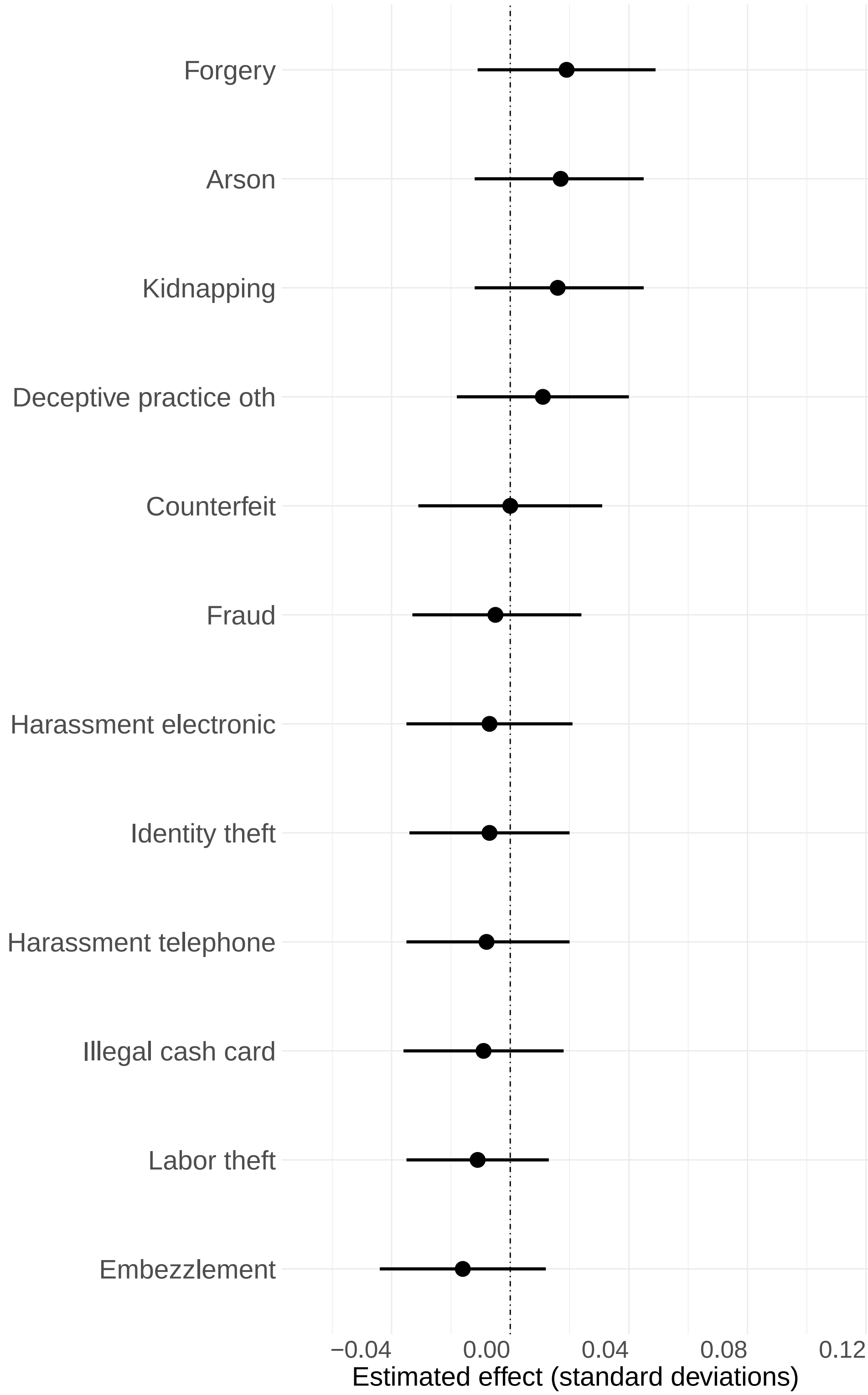

Figure 10: Placebo test: gangs do not cause crimes of deception, arson, or kidnapping

Note: This figure presents standardized regression coefficients from model (2) for crimes that have no obvious connection to gang activity. The bars represent two standard error intervals. There are over 500 distinct description codes in the raw data, which I aggregate into ≈ 138 categories by fixing spelling errors and lumping together crimes that are very similar. With a small number of exceptions (e.g., there were three description codes that fell under deceptive practice that involved theft), I present all description categories that fell under the IUCR codes of arson, kidnapping, and deceptive practice.

Table 1: Gang members are disproportionately young, male, and involved in violence

Characteristic	Gang Member, N = 64,947	Not Gang Member, N = 333,737
Age at Latest Arrest		
less than 20	19%	16%
20-30	42%	34%
30-40	22%	22%
40+	18%	27%
Sex		
F	3.1%	28%
M	97%	72%
X	<0.1%	<0.1%
Number of Times Victimized by Shooting		
1	7.1%	0.7%
2+	0.8%	<0.1%
Number of Times Victimized by Battery		
1	7.8%	1.9%
2+	1.2%	0.2%
Number of Times Arrested for Violence		
1	19%	5.6%
2+	6.2%	0.6%
Number of Times Arrested for Narcotics		
1	29%	13%
2+	23%	3.1%
Number of Times Arrested for Weapons Use		
1	9.4%	1.6%
2+	1.0%	<0.1%

Note: This table presents summary statistics for suspected gang members and non-gang members as calculated from the universe of individuals over 18 who were arrested in Chicago between the years of 2012 and 2016.

Table 2: Average effect of gangs on crime for major IUCR crime categories

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	0.133	0.221	0.058	-0.038
	(0.042)	(0.052)	(0.027)	(0.012)
Control group mean	2.201	1.195	0.111	0.478
	Assault	**Trespassing**	**Homicide**	**Weapons violation**
Post gang occupation	-0.009	0.075	0	0.012
	(0.018)	(0.03)	(0.002)	(0.006)
Control group mean	0.754	0.38	0.017	0.126

Note: This table presents estimates from model (2) and represents the average causal effect of gang occupation across the indicated IUCR crime categories. All regressions contained 7,524 census blocks, each observed for ≈11 years, for a total of 79,702 observations. Standard errors were calculated by clustering at the match-pair level as suggested in Abadie and Spiess (2019); however, I obtain similar levels of precision when clustering at higher levels of aggregation (e.g., the census tract). See Appendix D.2 for more detail.

Table 3: Results are robust to a variety of threats to identification

	(1)	(2)	(3)	(4)	(5)	(6)
Assault	-0.009	-0.017	-0.01	-0.001	-0.01	0.002
	(0.018)	(0.017)	(0.018)	(0.019)	(0.018)	(0.019)
Battery	0.133	0.097	0.126	0.166	0.125	0.152
	(0.042)	(0.037)	(0.041)	(0.043)	(0.042)	(0.043)
Trespassing	0.075	0.05	0.076	0.072	0.061	0.068
	(0.03)	(0.023)	(0.03)	(0.03)	(0.028)	(0.028)
Homicide	0	0.001	0	0	0	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Narcotics	0.221	0.175	0.226	0.237	0.195	0.261
	(0.052)	(0.046)	(0.05)	(0.054)	(0.049)	(0.059)
Prostitution	0.058	0.056	0.054	0.055	0.056	0.065
	(0.027)	(0.026)	(0.026)	(0.028)	(0.026)	(0.031)
Robbery	-0.038	-0.042	-0.04	-0.037	-0.039	-0.037
	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	(0.013)
Weapons Violation	0.012	0.012	0.012	0.009	0.012	0.014
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	79,702	79,702	79,702	79,702	79,702	79,702
Observations (blocks)	7,524	7,524	7,524	7,524	7,524	7,524
Baseline crime trends	no	yes	no	no	no	no
Demographic trends	no	no	yes	no	no	no
Gang database controls	no	no	no	yes	no	no
Public housing demo controls	no	no	no	no	yes	no
Police station trends	no	no	no	no	no	yes

Note: This table presents results from specification (2) except I have included different sets of controls meant to address specific threats to identification. Column (1) contains the main results from Table 2 as a baseline. Column (2) controls for a linear trend interacted with baseline measures of every IUCR crime type. Column (3) controls for a linear trend interacted in baseline census demographic variables. Column (4) adds controls for the number of individuals the local police beat adds to the gang database. Column (5) controls for flexible measures of exposure to public housing demolitions. Column (6) controls for a linear trend interacted with the distance of each block to every police station in the city.

TABLE 4: Results are similar aggregated to the census tract level

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	1.918	3.873	0.556	-2.25
	(0.696)	(0.947)	(0.241)	(0.201)
Control group mean	80.71	45.09	3.29	16.92

	Assault	Trespassing	Homicide	Weapons violation
Post gang occupation	0.213	1.2	0.097	0.071
	(0.203)	(0.355)	(0.014)	(0.064)
Control group mean	28.27	14.08	0.6	4.63

Note: This table presents estimates from model (2) except the data have been aggregated to the census tract level. See Section 6.2 for details regarding the aggregation procedure. Standard errors are clustered at the census tract level.

TABLE 5: Dif-in-dif estimates of the impact of gang occupation on ACS outcomes

	(1)	(2)	Control mean
Median hh income	-1,755.4	-1,866.8	59,711.1
	(901.2)	(965.8)	
Median house value	-9,035.7	-8,436.9	261,123.9
	(4158.2)	(4478.3)	
Per capita income	-614.4	-585.6	33,542.4
	(635.6)	(630.9)	
Total population	-106.4	-109.2	4,101.2
	(67.5)	(64.7)	
Observations	1032	1032	
Census tract fixed effects	no	yes	
Match-pair x post-period fixed effects	no	yes	

Note: The table presents results from model (3) applied to the matched and stacked sample as described in Section 6.2. Standard errors are clustered at the match-pair level as suggested in Abadie and Spiess (2019).

TABLE 6: Gangs cause larger increases in crime in more competitive environments

	Battery				Narcotics			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post	0.133	0.237			0.221	0.312		
	(0.042)	(0.045)			(0.052)	(0.049)		
Post x gang-sets		0.014				0.022		
		(0.007)				(0.009)		
Post x members		-0.002				-0.003		
		(0.001)				(0.001)		
One gang			0.113				0.105	
			(0.038)				(0.041)	
Two+ gangs			0.182				0.255	
			(0.076)				(0.104)	
Near				0.043				0.087
				(0.057)				(0.069)
Border				0.111				0.15
				(0.047)				(0.052)
Deep				0.085				-0.059
				(0.071)				(0.08)

Note: The table presents estimates from model (2) but using different versions of the treatment variable that allow the causal effect of gang entry to vary with the competitiveness of the market. "Post" is the post-entry indictor from the baseline model. "Gang sets" counts the number of distinct groups present within each gang for the 1982 birth cohort. "Members" counts the number of gang members in the 1982 birth cohort. "One gang" takes a value of one when there is a gang within a quarter mile of the block. "Two+ gangs" takes a value of one when there is more than one gang within a quarter mile of the block. "Near" takes a value of one when a block is close to gang territory but is not on the border. "Border" takes a value of one when a block is on the border of gang territory. "Deep" takes a value of one when a gang is inside gang territory but is not on the border. Standard errors are clustered at the match-pair level.

Table 7: The supply of crime does not appear to vary with the elasticity of law enforcement

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	0.135	0.215	0.058	-0.037
	(0.043)	(0.054)	(0.028)	(0.013)
Post x distance to CPD (miles)	-0.016	0.063	0.001	-0.015
	(0.042)	(0.045)	(0.027)	(0.012)

	Assault	Trespassing	Homicide	Weapons violation
Post gang occupation	-0.01	0.078	0.001	0.014
	(0.018)	(0.032)	(0.002)	(0.006)
Post x distance to CPD (miles)	0.01	-0.028	-0.001	-0.011
	(0.018)	(0.03)	(0.002)	(0.006)

Note: The table presents estimates from model (2) but the post-entry indicator is interacted with distance to the nearest police station. If police officers find it more costly to patrol more remote locations, then the elasticity of law enforcement with respect to the supply of crime should be smaller at further distances. Standard errors are clustered at the match-pair level.

Intended for Online Publication Only

A Data Appendix

A.1 Chicago Police Department data on gang boundaries

The CPD provided me with a shapefile for each year from 2004 to 2012 and 2014 to 2017. Subsequently, another research team sent a FOIA request to obtain the 2018 gang map, and this team shared that map with me. The officer who provided me with the data informed me that the gang boundaries did not change from 2012 to 2013; hence there was no need for them to provide a separate shapefile for that year.

I read in the shapefiles using the R "Simple Features" package (Pebesma, 2018). This results in 14 data frames of gang geospatial information. For 2013, I duplicate the 2012 data frame. Within each geospatial data frame, an observation corresponds to a gang, and for each gang, there is an associated "geometry" column that contains the boundaries of their territory within the city during the indicated year.

I manually fix several naming inconsistencies that appear across years in the data. For example, in some years, I view a gang named "ylo cobras," and in other years I view a gang named "Young Latin Organization Cobras." I assume these are the same gangs. There are occasionally geometries in the CPD data that do not have a gang name associated with them. I drop these from the data. There are also some geometries labeled with two gang names. For example, there is a geometry in one year labeled "mix unknown and traveling vice lords." In these cases, I assume the polygon was occupied by both gangs.

A.2 Incident-level geospatial crime data

Incident-level geospatial crime data are publicly available on the Chicago open data portal. From January 1, 2001 to September 9, 2020,[36] there are 7,195,964 distinct incidents in the data. Over 99% of the incidents contained in the data have an associated latitude and longitude. I drop the incidents that are missing these coordinates.

To determine the block where these crimes occurred, I intersect the latitudes and longitudes associated to the incidents with the year 2000 census block polygons contained in a shapefile publicly available on the Chicago open data portal. There are 16,335 incidents that do not intersect with any census block. I drop these incidents.

[36]Which is the most recent date that I downloaded the data

From there, I aggregate the data in two ways. First, I aggregate the incident data by block, year, and IUCR crime type. Second, I perform a similar aggregation using the strings contained in the "description" field of the raw incident data. The raw data contain 528 unique strings in the description field. I go through these strings by hand and recode them into 139 unique strings. Most of this recoding involves fixing spelling mistakes/inconsistencies, combining descriptions together that were similar, or lumping extremely infrequent incidents together. For example, the following are three distinct descriptions in the raw data that all fall under the IUCR category of motor vehicle theft: "CYCLE, SCOOTER, BIKE W-VIN," "CYCLE, SCOOTER, BIKE WITH VIN," and "SCOOTER, BIKE NO VIN." In this case, I combine all three description codes into a single category. There is also a description called "ATTEMPT: CYCLE, SCOOTER, BIKE NO VIN," which only occurs seven times in the incident data. I combine this rare incident with the other three categories referenced previously.

A.3 Data from Kirby et al. (2006)

Within the 2006 edition of Kirby et al. (2006), there is a section that contains short narrative histories for 56 gangs. For the gangs that appear both in the CPD data and Kirby et al. (2006), I manually transcribed information related to the primary race/ethnicity of the gang and any history of conflict or rivalry. When a gang was missing from Kirby et al. (2006), I supplemented information on the race/ethnicity of the gang from the website chicagoganghistory.com. In conversations with the owner of this website, I learned that he collects information on Chicago street gangs primarily from court records and urban folklore as passed on to him by email from former gang members and residents of gang-occupied neighborhoods.

There are a small number of gangs that have names evocative of a larger gang but do not appear as independent gangs in either the "Gang Book" or on chicagoganghistory.com. These gangs also often appear as sharing territory with the larger gang. For example, there are gangs named "Traveling Vice Lords" and "Conservative Vice Lords," which overlap geometries in the CPD data with a polygon simply labeled "Vice Lords." Thus when this naming convention occurs and I cannot find evidence in Kirby et al. (2006) or on chicagoganghistory.com that these are, in fact, two separate entities, I assume that the smaller gang is part of the larger gang and merge the two geometries.

Finally, there are a small number of gangs (11) that do not have names, suggesting they belong to larger factions, and of which I can find no evidence of their existence in either Kirby et al. (2006) or on chicagoganghistory.com. For these gangs, I merge them into a single geometry. In an average year, these gangs account for just under 1% of gang territory in the city.

A.4 Graffiti data

The city of Chicago operates a program where city residents may call to report graffiti in their neighborhood, and the city will send out a team from the Department of Streets and Sanitation to remove the graffiti. The city makes information on these calls for service available on the Chicago open data portal website. The data contain approximately one million observations spanning 2011 to 2018. Each observation is a removal request and contains variables describing the latitude and longitude where the graffiti is reportedly located. I drop 180 removal requests that were listed as "open" or "open-dup." I drop 607 removal requests that were missing the associated latitude or longitude. I also find 2,161 removal requests that were reported to have occurred prior to 2011. However, the dates were often non-sensical (the earliest year was 1926), and to my knowledge the database did not exist prior to 2011, so I drop these removal requests as well. From there, I aggregate the remaining removal requests to the census block level using the associated shapefiles. I could not geolocate 376 of the observations to any census block in the city, so I drop them as well.

B Additional descriptive statistics

B.1 Gang size distribution by area

In Section 4 of the main text, I use estimated gang membership to make the claim "Chicago has a large number of small gangs." In Figure A.1, I show that this same conclusion holds if I measure gang size using the geographic area occupied by individual gangs in 2004 rather than estimated membership.

[Figure A.1 about here.]

B.2 Violence and gang territory controlling for population

In Section 4 of the main text I document interesting patterns relating to the mean and variance of violence in border regions. Here I show that this does not appear to be a pure population effect. Figure A.2 is identical to Figure 4 from the main text, except I have replaced the independent variable with residuals from a regression of violent crime in 2004 on total population. To allow the relationship between crime and total population to vary non-parametrically, I discretize the total population variable into 100 total population indicators that each contain 1% of the data. I prefer residualizing to adjust for population at the block level over scaling the dependent variable by population because there are several blocks having zero population and that also exhibit violent crime.

[Figure A.2 about here.]

B.3 Additional event study plots

In Figure A.3, I present event study plots for the IUCR crime categories that are explored in the main text but were not included in Figure 8.

[Figure A.3 about here.]

B.4 Formal test for pre-trends

In this section I present formal tests for pre-trends. Specifically, I take every IUCR crime category that exhibits enough variation to identify the relevant event study coefficients and estimate model (1).[37] For each IUCR crime category, I then test the joint hypothesis that all event study coefficients in the pre-period are equal to zero. Figure A.4 displays the distribution of p-values generated by this exercise. Among the eight crimes explored in the main text, the smallest p-value is 0.16 and belongs to narcotics violations.

[Figure A.4 about here.]

C Details of matching procedure

To form the matched comparison group, I first estimate the following linear probability model:

$$d_b = \alpha + \beta X_b + \epsilon_b, \tag{5}$$

where d_b takes a value of one if block b ever experiences an occupation event between 2004 and 2018 and X_b is a vector of covariates that includes crime counts in 2001 and 2002 for every IUCR crime category contained in the incident data in those years as well as year 2000 census

[37] Of the 32 official IUCR crime categories, there are four crimes that are rare enough that I am unable to identify all of the event study coefficients on the matched sample.

variables that count population by race and ethnicity and the number of owner-occupied, renter-occupied, and vacant houses; and variables that describe the total length and number of city streets in the block.

I then use this propensity score to match without replacement between the blocks that do and do not experience an occupation event. I do this by randomly choosing a treated block. Then I restrict the pool of potential control blocks to those that have not already been matched and that do not appear in the same police district. The second condition is to force some geographic distance between treatment and control groups and thus help guard against the possibility of spillovers that could otherwise contaminate the comparison unit. I repeat this procedure until all units are matched.

Figure A.5 displays the support of the propensity scores for the treated blocks and the matched comparison blocks. From the figure, we can see that the matched comparison blocks and the treated blocks share a common support.

[Figure A.5 about here.]

Table A.1 provides summary statistics for the treated blocks and the matched comparison blocks for crime in the year 2002 and for demographic variables as measured in the year 2000 census.

[Table A.1 about here.]

D Additional robustness checks

D.1 Robustness to alternative matching procedures

In this section, I present three robustness checks meant to probe the degree to which the main results in the paper depend on the particulars of the matching procedure. In Table A.2, I present results from models that are identical to Table 2, except that I form that matched comparison group using Mahalanobis distance in place of the propensity score. In Table A.3, I present results from models that are identical to Table 2, except that I only use the baseline crime variables in the matching step. In Table A.4, I present results from models that are identical to Table 2, except that I omit the crime variables and only use demographic and road variables in the matching step. In all cases, the results are broadly similar to those found in Table 2.

[Table A.2 about here.]

[Table A.3 about here.]

[Table A.4 about here.]

D.2 Clustering standard errors at higher levels of aggregation

In this section, I present results that are identical to those contained in Table 2 except that I cluster the standard errors at the census tract level (Table A.5) and at the community area level (Table A.6).

[Table A.5 about here.]

[Table A.6 about here.]

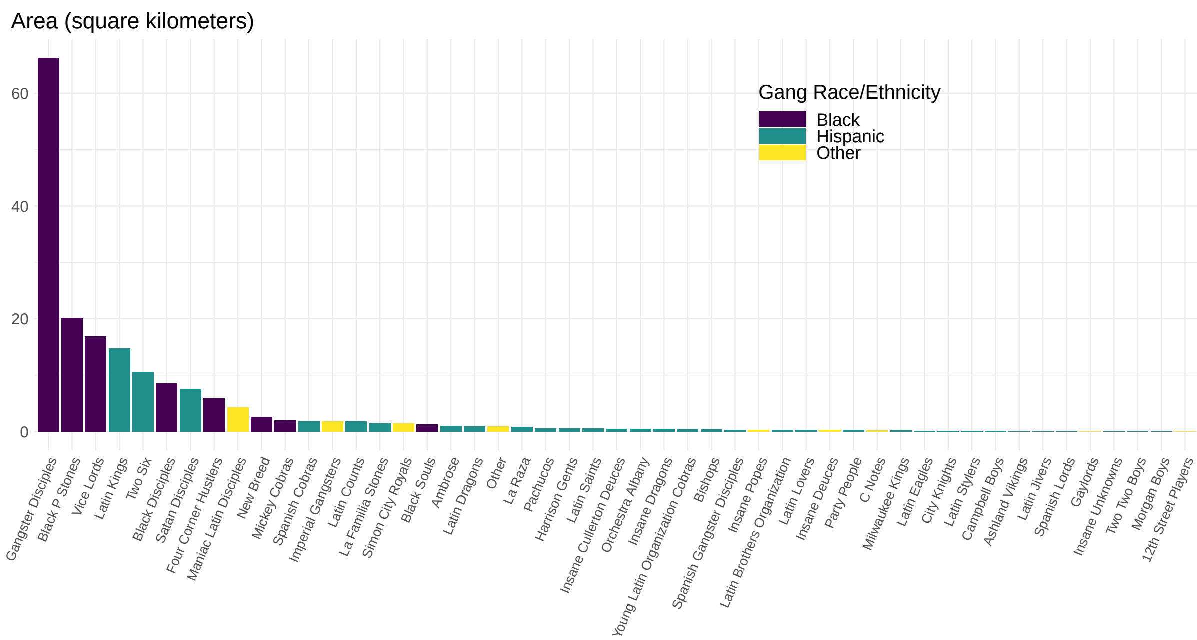

FIGURE A.1: Chicago has a large number of small gangs (as measured by area)

Note: This figure is identical to Figure 1 from the main text, except that I plot area (measured in square kilometers) on the y-axis instead of membership. Thus the broad conclusion from Figure 1 continues to hold if gang size is measured using area instead of membership.

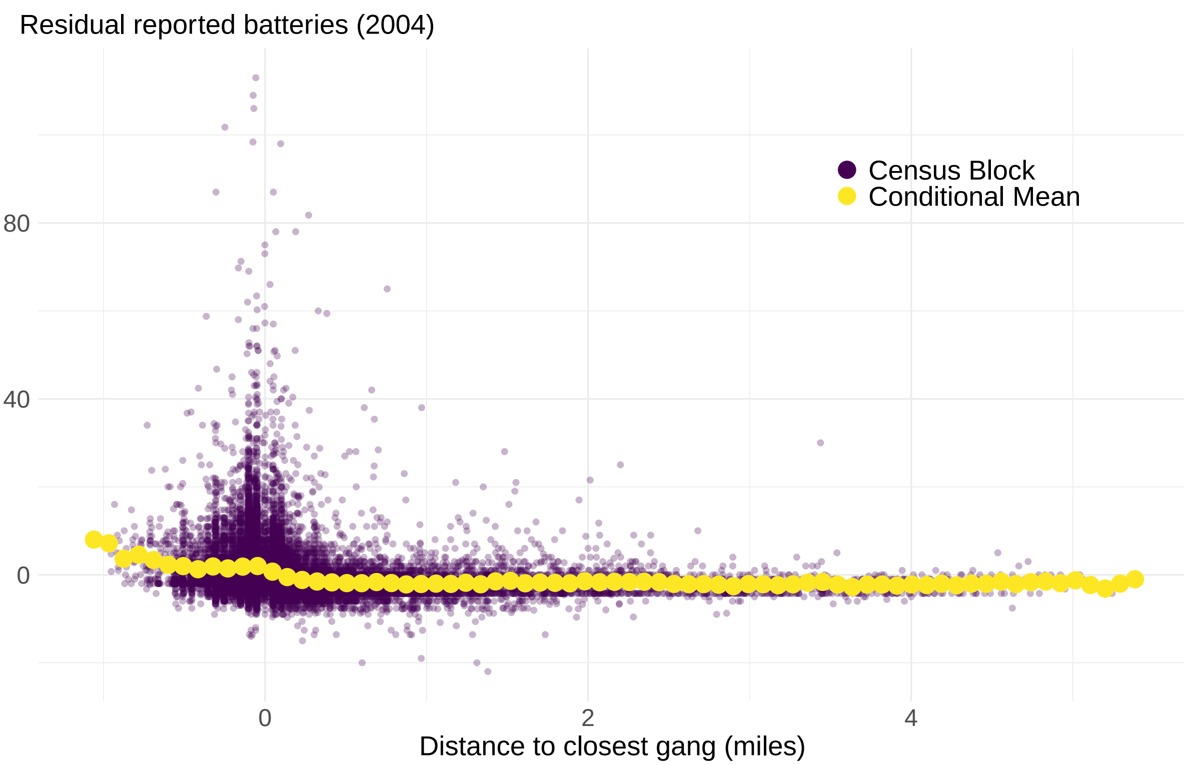

Figure A.2: Violence and gang borders (controlling for population)

Note: This figure plots residual counts of batteries against distance to the nearest gang border. Residuals are constructed by regressing batteries in 2004 on a semi-parametric representation of total population as measured in the year 2000 census. Small points with dark shading denote census blocks. Closest distance is defined as the minimum over all gangs of the minimum distance between the census block centroid and the individual gang border. Negative values of distance denote blocks that are inside gang territory. Large, lightly shaded points denote conditional means within 70 equal-length distance bins.

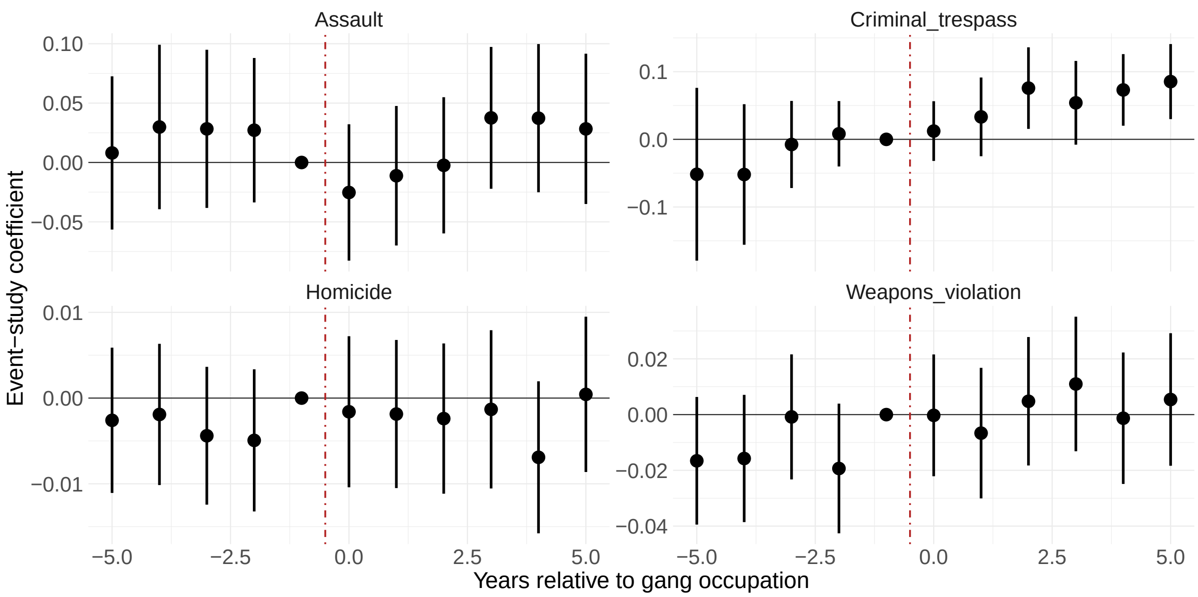

FIGURE A.3: Additional event study plots

Note: This figure displays event study coefficients from regression (1) for four outcomes of interest that were not displayed in the main text: assault, criminal trespassing, homicide, and weapons violations.

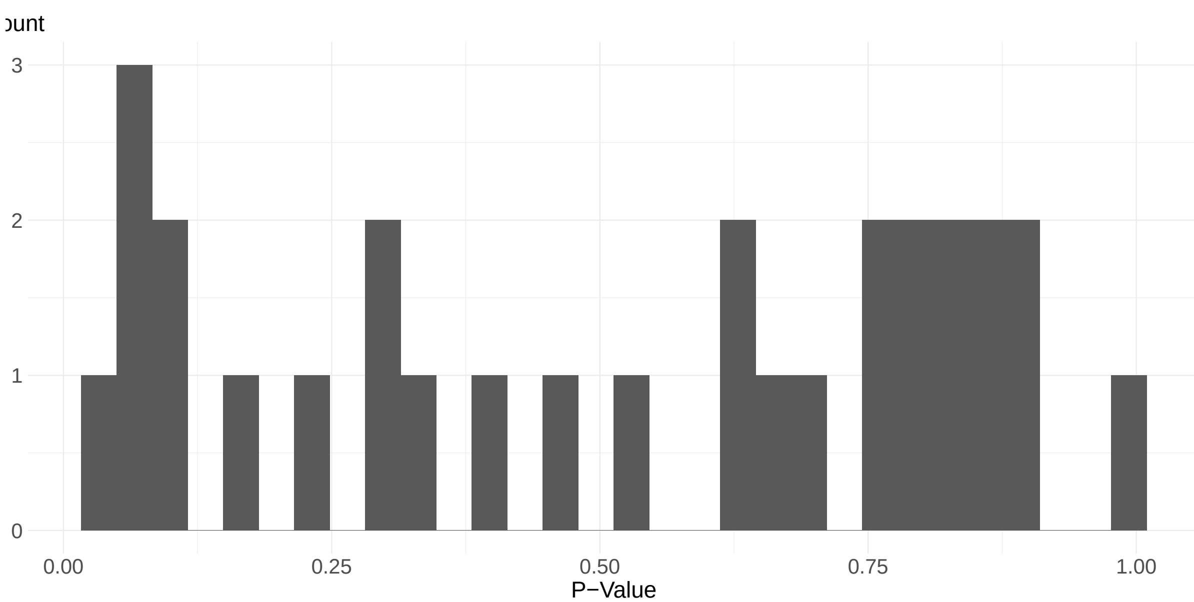

Figure A.4: Distribution of p-values for pre-trend tests conducted on every IUCR crime category

Note: This figure plots a histogram of p-values from formal tests for pre-trends applied to every IUCR crime category contained in the data.

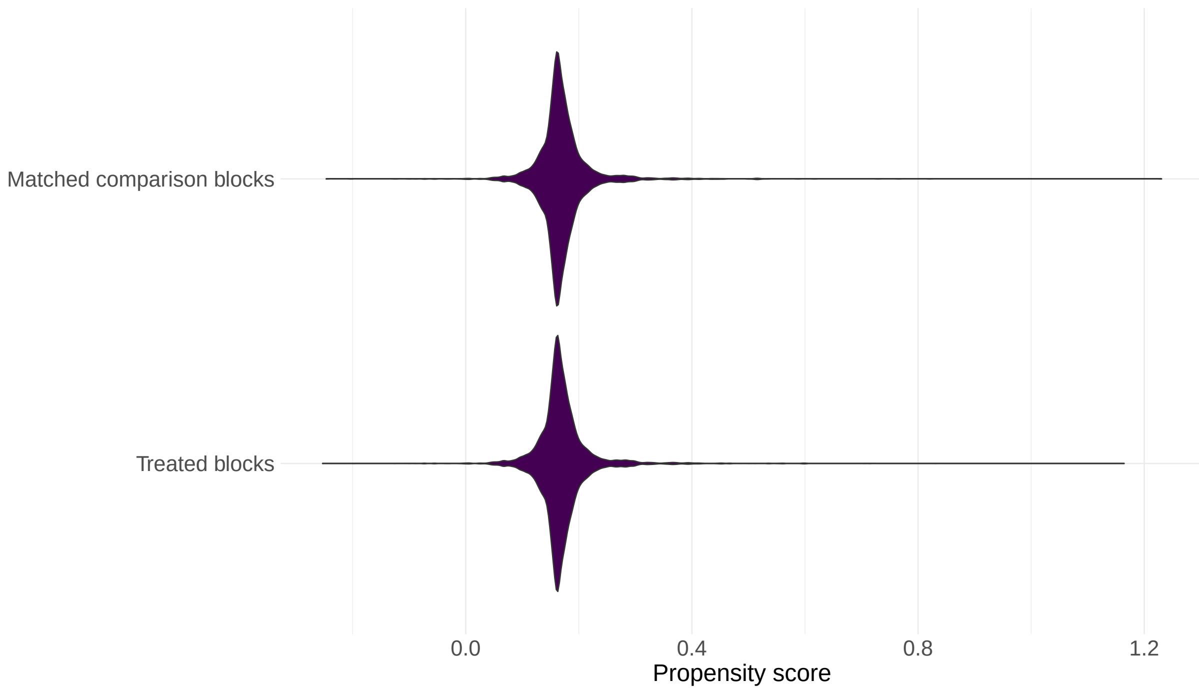

Figure A.5: Common support

Note: This figure displays the densities of the estimated propensity scores for the treated and matched comparison groups.

Table A.1: Covariate balance

Characteristic	Treated blocks, N = 3,762	Matched comparison blocks, N = 3,762
Arson	0.0287 (0.0029)	0.0303 (0.0029)
Assault	0.96 (0.03)	0.99 (0.03)
Battery	2.8 (0.1)	3.0 (0.1)
Burglary	0.97 (0.02)	0.99 (0.02)
Criminal damage	2.04 (0.04)	2.11 (0.04)
Criminal sexual assault	0.0550 (0.0040)	0.0569 (0.0045)
Criminal trespass	0.42 (0.04)	0.57 (0.07)
Deceptive practice	0.42 (0.03)	0.46 (0.02)
Gambling	0.0268 (0.0032)	0.0295 (0.0035)
Homicide	0.0138 (0.0021)	0.0173 (0.0024)
Interference with public officer	0.0096 (0.0016)	0.0117 (0.0021)
Intimidation	0.0053 (0.0012)	0.0064 (0.0014)
Kidnapping	0.0237 (0.0027)	0.0237 (0.0028)
Liquor law violation	0.0391 (0.0039)	0.0385 (0.0037)
Motor vehicle theft	0.93 (0.04)	0.94 (0.03)
Narcotics	1.17 (0.11)	1.45 (0.09)
Obscenity	0.0005 (0.0004)	0.0003 (0.0003)
Offense involving children	0.08 (0.01)	0.07 (0.00)
Other narcotic violation	0.0003 (0.0003)	0.0003 (0.0003)
Other offense	1.16 (0.03)	1.15 (0.03)
Prostitution	0.1956 (0.0359)	0.2310 (0.0386)
Public indecency	0.0000 (0.0000)	0.0003 (0.0003)
Public peace violation	0.06 (0.01)	0.06 (0.00)
Ritualism	0.0000 (0.0000)	0.0000 (0.0000)
Robbery	0.60 (0.02)	0.60 (0.02)
Sex offense	0.07 (0.01)	0.08 (0.00)
Stalking	0.0072 (0.0014)	0.0077 (0.0015)
Theft	3.3 (0.2)	3.3 (0.1)
Weapons violation	0.10 (0.01)	0.12 (0.01)
White	50 (1)	50 (1)
Black	38 (1)	38 (1)
Amer ind	0.36 (0.02)	0.35 (0.02)
Asian	9 (1)	8 (1)
Hawaiin	0.0683 (0.0089)	0.0579 (0.0072)
Other	13 (0)	13 (1)
Multi race	4 (0)	4 (0)
Owner occ	19 (0)	19 (1)
Renter occ	23 (1)	23 (1)
Vacant	3.2 (0.2)	3.1 (0.1)
Street num	2.22 (0.03)	2.26 (0.03)
Total length	1,888 (19)	1,946 (27)

Note: This figure displays averages for the treated blocks and the matched comparison blocks for crime in the year 2002, demographic variables as measured in the year 2000 census, and for variables related to the road network. Standard errors are displayed in parentheses.

Table A.2: Average effect of gangs on crime using Mahalanobis matching

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	0.201	0.372	0.063	-0.027
	(0.047)	(0.071)	(0.024)	(0.013)
Control group mean	2.356	1.42	0.105	0.507

	Assault	Trespassing	Homicide	Weapons violation
Post gang occupation	0.011	0.101	0.003	0.013
	(0.019)	(0.047)	(0.002)	(0.006)
Control group mean	0.795	0.422	0.02	0.141

Note: In this table, I present results from models that are identical to Table 2, except that I form that matched comparison groups using Mahalanobis distance in place of the propensity score.

Table A.3: Average effect of gangs on crime matching only on baseline crime

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	0.099	0.164	0.05	-0.044
	(0.041)	(0.054)	(0.023)	(0.013)
Control group mean	2.115	1.083	0.09	0.443
	Assault	**Trespassing**	**Homicide**	**Weapons violation**
Post gang occupation	-0.001	0.03	0.003	0.013
	(0.017)	(0.023)	(0.002)	(0.006)
Control group mean	0.725	0.312	0.016	0.124

Note: In this table, I present results from models that are identical to Table 2, except that I only use the baseline crime variables in the matching step.

Table A.4: Average effect of gangs on crime matching only on demographics

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	0.422	0.544	0.164	0.016
	(0.053)	(0.08)	(0.041)	(0.014)
Control group mean	3.07	1.805	0.21	0.675

	Assault	Trespassing	Homicide	Weapons violation
Post gang occupation	0.065	0.12	0.005	0.018
	(0.021)	(0.026)	(0.002)	(0.006)
Control group mean	1.063	0.501	0.022	0.175

Note: In this table, I present results from models that are identical to Table 2, except that I only use the demographic and road variables in the matching step.

Table A.5: Average effect of gangs on crime, standard errors clustered at the census tract

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	0.133	0.221	0.058	-0.038
	(0.033)	(0.042)	(0.021)	(0.011)
Control group mean	2.201	1.195	0.111	0.478
	Assault	**Trespassing**	**Homicide**	**Weapons violation**
Post gang occupation	-0.009	0.075	0	0.012
	(0.013)	(0.02)	(0.001)	(0.004)
Control group mean	0.754	0.38	0.017	0.126

Note: In this table, I present results from models that are identical to Table 2, except that I cluster the standard errors at the census tract level.

Table A.6: Average effect of gangs on crime, standard errors clustered at the community area

	Battery	Narcotics	Prostitution	Robbery
Post gang occupation	0.133	0.221	0.058	-0.038
	(0.04)	(0.051)	(0.021)	(0.013)
Control group mean	2.201	1.195	0.111	0.478
	Assault	Trespassing	Homicide	Weapons violation
Post gang occupation	-0.009	0.075	0	0.012
	(0.017)	(0.019)	(0.001)	(0.005)
Control group mean	0.754	0.38	0.017	0.126

Note: In this table, I present results from models that are identical to Table 2, except that I cluster the standard errors at the community area level.