The Long-term Effects of Structural Discrimination on Public Safety: The 1930s Redlining Maps*

David Mitre-Becerril†
University of Connecticut

August 18, 2024

Abstract

Enduring disinvestment can lead to different growth trajectories and the clustering of distressed, blighted urban areas. Do these differential trajectories explain present-day public safety? This study examines the long-term public safety impacts of the residential security maps, the once-legal 1930s racially discriminatory maps used to determine the real estate market "risk" of lower property values, limiting residential loans to racial-minority creditworthy individuals. This research collects incident-level data from 48 police departments and leverages within-city variation by comparing areas near different color-grade boundaries around a small bandwidth in a spatial regression discontinuity design. Sociodemographic differences existed before the implementation of the redlining maps. There are significant increases in present-day violent, property, low-level, and weapon-related offenses in redlined areas, even after controlling for historical differences. There is some evidence that present-day law enforcement, measured via arrests, police stops, and calls for service, was also impacted. The results suggest that the redlining maps exacerbated neighborhood disadvantages, preventing safer neighborhoods in the long run, though its impact may be smaller than previous evidence has suggested.

Keywords: redlining, crime, community investments, discrimination, regression discontinuity

*I thank Aaron Chalfin, John M. MacDonald, Greg Ridgeway, Anthony Braga, Lindsey Bullinger, and seminar participants from the University of Pennsylvania, the Academy of Criminal Justice Sciences Annual Meeting, the Criminology Consortium Annual Meeting, and the Association for Public Policy and Management for helpful comments and suggestions. Uyen Le provided a great research assistant. This research was partially funded by the University of Connecticut College of Liberal Arts and Sciences Diversity, Equity, and Inclusion Initiatives.

†School of Public Policy, david.mitre@uconn.edu

1 Introduction

Institutions can promote economic development and reduce inequalities by targeting disadvantaged places (Bartik, 2020; Ladd, 1994; Neumark and Simpson, 2015). However, they can also cause the opposite effect by purposely withholding the enjoyment of benefits to specific groups. In the US, there is a longstanding urban economics history of economic disadvantage and inequality caused by government-sanctioned policies creating adverse effects on racial-minority groups (Hillier, 2003; Jones-Correa, 2000; Massey, 2015; Von Hoffman, 2008; Wilson, 2008). A commonality among them is the impact of housing policy and property development on the growth and clustering of segregated neighborhoods. Judicial decisions and new legislation have struck down several of those policies. Still, it is an open question whether they affected communities and individuals.

This research focuses on understanding the long-run effects of one of those policies, redlining, on criminal and police activity. Redlining was coined for the racially discriminatory practice used by the Home Owners Loan Corporation (HOLC) and the Federal Housing Authority (FHA) in the 1930s to determine the real estate market "risk" of lower property values and limiting mortgage loans to racial-ethnic minority creditworthy individuals. Appraisers explicitly used race and ethnicity to develop the four-color-grade residential security maps. The maps may have aligned the lender's and realtors' expectations about the neighborhoods deemed "obsolete" and "undesirable", depressing housing values due to the presence of racial-minority residents and poor housing conditions.

Residential segregation was not new in the 1930s, nor was it illegal back then, but the relevance of these maps was its application to an unprecedented level (Faber, 2021). It took over a decade to remove the explicit reference to racial motivations from the FHA underwriting manual precluding individuals from obtaining a mortgage (Jones-Correa, 2000), and almost three decades to legally ban discrimination in the housing acquisition market (Massey, 2015). Even after removing legal barriers towards housing discrimination, recent literature has found that the residential security maps caused long-lasting negative impacts on individuals and neighborhoods (Aaronson et al., 2021a,b; Anders, 2023; Appel and Nickerson, 2016; Benns et al., 2020; Faber, 2020, 2021; Jacoby et al., 2018; Lyons et al., 2023; Mitchell and Chihaya, 2022; Nardone et al., 2020; Poulson et al., 2021).

There are several mechanisms by which redlining could have affected public safety. Material deprivation has been an enduring topic in the crime literature (Merton, 1938; Wright, 1893). Limited localized investments can inhibit the neighborhood's ability to maintain vitality and reduce the opportunity cost of crime (Becker, 1968; Clarke and Cornish, 1985). Incivilities can induce fear among residents, reduce social interactions, and weaken social controls, leading to serious criminal activity (Kelling et al., 1982). Hindering residents' social interactions also prevents realizing shared values and using social controls and formal institutions to prevent antisocial behaviors (Bursik Jr, 1988; Sampson and Groves, 1989). The impact of neighborhood disadvantage on police activity is more nuanced, mainly as race has been longstanding connected to poverty, joblessness, and social isolation (Wilson, 2003, 2008). The local racial composition seems to matter in determining police enforcement (Braga et al., 2019; Brunson and Weitzer, 2009) as police officers may use race as a predictor of suspicious behavior, increasing police surveillance of minority groups (Alpert et al., 2005). Enhanced policing can benefit Black communities by reducing homicides, but they disproportionately face more low-level quality-of-life arrests (Chalfin et al., 2022). Moreover, law enforcement seems responsive to real estate and urban development investments (Beck, 2020; Laniyonu, 2018).

The HOLC color-grade areas were not assigned randomly, so comparing them faces the risk of confounding the impacts on public safety. Previous research has overcome this concern by making pre-post-HOLC maps comparisons between areas in a difference-in-differences with propensity score matching approach (Aaronson et al., 2021b,a; Faber, 2020). These studies leverage historical data from over a century ago to measure long-term outcomes. Some outcomes, like criminal activity, lack pre-intervention data to measure pre-post changes. Accordingly, such studies have contrasted cross-sectional data among large areas, usually census tracts (Benns et al., 2020; Faber, 2021; Jacoby et al., 2018; Mitchell and Chihaya, 2022; Poulson et al., 2021). Others have compared only the areas closer to the borders to reduce confounding factors (Anders, 2023; Appel and Nickerson, 2016). These last studies use a spatial regression discontinuity design by leveraging the quasi-random assignment of the borders. The underlying idea is that map appraisals may have agreed on how they should have classified large areas, but they were likely uncertain about where exactly to draw the borders. By creating discretionary color-grade boundaries where they did not exist, the residential security maps likely aligned the stakeholders' beliefs and created a sharp discontinuity in investment decisions. However, recent research (Aaronson et al., 2021a; Fishback et al., 2023), along with this manuscript, highlights existing pre-intervention significant differences between adjacent color-grade borders. Not accounting for such pre-existing disadvantaged inequalities tends to overestimate the long-term effects of redlining on crime.

This research collected and georeferenced incident-level public safety administrative records between 2014 and 2019 from 48 US cities. Using a spatial regression discontinuity, the results reveal significant crime increases across all categories: violent, property, low-level, and weapon-related offenses are 10 to 32 percent larger in the downgraded areas than their similarly situated better-grade counterparts. Once accounting for pre-intervention (1900-1930) sociodemographics controls, the long-term impacts on crime are still statistically significant but are reduced by one-third: violent, property, low-level, and dangerous weapon crimes increased around six to 15 percent in downgraded places. The redlining impacts on police arrests tend to be null findings, except the dangerous weapon offenses, with an increase of 22 percent after controlling for historical sociodemographic differences. Police stops show a marginally significant increase of around 16 percent, while calls for service experience a decrease of about 10 percent. Both outcomes are imprecisely estimated. The results suggest that redlining increased crime across the board but only significantly increased police enforcement for dangerous weapons, which usually is a by-product of police proactivity. Redlining may have decreased calls for service, a marker for citizens' cooperation and civic engagement. In short, the evidence highlights the long-run adverse effects of structural discrimination on public safety.

Finally, this research makes several contributions. First, it provides a quasi-experimental design (regression discontinuity) that more precisely estimates the impact of redlining on crime and uses pre-intervention (1900-1930) data to account for historical differences. Second, it uses geographical granular information (a hexagonal area of 0.025 squared kilometers or approximately a city block) to measure changes near a small area at the HOLC grade border. Next, it relies on information from 48 cities, considerably expanding the results' external validity and statistical power to detect small impacts. Fourth, it provides detailed crime information from six Uniform Crime Reporting Part I offenses and the aggregated offenses from Part II. Importantly, it is the first study to present evidence on law enforcement activity. Specifically, it disaggregates police arrests into the same Part I and Part II categories and measures police stops and 911 emergency calls for service. Law enforcement arrests, stops, and calls for service allow for measuring police proactivity and citizens' cooperation (Brantingham and Uchida, 2021; Xie and Baumer, 2019).

The remaining article is organized as follows. Section 2 provides a brief history of residential security maps and the literature review. Sections 3 and 4 explain the data and empirical strategy. Section 5 presents the results, and Section 6 concludes.

2 Background

2.1 Security residential maps

The 1929 Great Depression decreased house prices and increased foreclosures to unprecedented levels. The housing market had one of the slowest recovery rates in the early 1930s (Fishback et al., 2011; Rose, 2011). As a result, policymakers enacted several laws to stabilize the mortgage industry and help current and future homeowners acquire and keep their properties. The Home Owners Loan Act of 1933 created the government-sponsored agency HOLC to provide emergency relief on home mortgage indebtedness. Between 1933 and 1935, HOLC's rescue phase focused on refinancing home mortgages with more generous terms so homeowners could save their properties: lower interest rates, more extended amortization periods (15 years), and fully amortized loans (lendee pays the principal and interest simultaneously). While these loan features are standard nowadays, that was not the case in the 1930s. HOLC outreach was effective as nearly 40 percent of the eligible homeowners applied for refinancing their mortgage, and almost half received it (Hillier, 2005; Michney and Winling, 2020).

Over the following 16 years (1935-1951), the consolidation phase managed and sold the acquired housing inventory held in the mortgages until HOLC's eventual liquidation. The agency also compiled data to improve the housing market. In 1935, HOLC's parent organization, the Federal Home Loan Bank Board (FHLBB), was mandated to survey all cities above 40,000 people to gather data about the local real estate conditions and market risks. It relied on HOLC to carry out this responsibility, formally known as the City Survey Program. Taking advantage of HOLC's staff in these locations and their connections to local realtors and lenders, in addition to hiring map consultants, the program resulted in residential security maps for 239 cities reflecting the residential desirability of the neighborhoods (Hillier, 2005; Michney and Winling, 2020). These maps used a four-color grade system, which Hillier (2005, p. 216-217) describes using the FHLBB records' quotes:[1] grade A (green) were "homogeneous", on-demand areas during "good times or bad"; grade B (blue) included places that had developed entirely, meaning that they were "still good, but not what the people are buying today who can afford a new one"; grade C (yellow) grouped the older, almost obsolete neighborhoods that had "infiltration of a lower grade population"; and grade D (red) had poor housing conditions, "undesirable population or an infiltration of it", and "represent those neighborhoods in which the things that are now taking place in the C neighborhoods, have already happened". The explicit discriminatory use of race and ethnicity in developing the residential security maps faced no legal obstacles.

Another relevant legislation to address the housing market problems was the 1934 National Housing Act, which created the FHA. Its purpose was to regulate interest rates and mortgage loans by collaborating with banks and other financial institutions. It offered insurance on home mortgages to guarantee loan repayment to the lenders in case of a default. These loans had to satisfy several criteria. They had to provide a fixed interest rate, a long-term plan, and be fully amortized (similar to the features provided by HOLC). In addition, the loans needed to meet the FHA underwriting and construction standards. These requirements

[1] A detailed online description of the grading system is available on Aaronson et al. (2021b) online Appendix.

standardized and revolutionized the mortgage lending industry and provided certainty to lenders about their long-term financial planning and better terms for new homeowners.[2] The FHA Underwriting and Valuation Procedure Manual had explicit racially discriminatory guidelines. For instance, Manual's section 233 stated that "[...]If a neighborhood is to retain stability, it is necessary that properties shall continue to be occupied by the same social and racial classes" (FHA, 1936, p. 198), so racial minority groups were discouraged from acquiring loans in predominantly White neighborhoods. Similarly, section 284 prohibited "the occupancy of properties except by the race for which they are intended" (FHA, 1936, p. 211); an explicit reference to racial covenants, a practice that was present in the real estate market since the early 1900s (Jones-Correa, 2000), which has also caused harmful long-term impacts on racial segregation and homeownership among Blacks (Sood et al., 2019).

Historians debate whether HOLC or FHA was the primary agency permeating the discriminatory lending practices across the mortgage industry. It is unclear which stakeholders and institutions had access to the appraisal maps and the extent to which they guided their lending decisions solely on these cartographic tools (Hillier, 2003; Jackson, 1980; Michney and Winling, 2020; Ryan, 2018; Woods, 2012). While the precise historical access and use of the residential security maps in the mid-20th century is still an open question, there is a consensus that the discriminatory practice of lending based on the neighborhood racial composition has caused long-lasting adverse effects.

2.2 Research on redlining and neighborhood conditions

A growing literature shows the long-lasting harmful effects of redlining on income, poverty, social mobility (Aaronson et al., 2021a), racial isolation (Faber, 2020), real estate ownership (Aaronson et al., 2021b; Appel and Nickerson, 2016; Faber, 2021; Hillier, 2003), health diseases, emergency visits (Nardone et al., 2020), intra-urban heat (Hoffman et al., 2020), and crime and violence (Anders, 2023; Benns et al., 2020; Jacoby et al., 2018; Lyons et al., 2023; Mitchell and Chihaya, 2022; Poulson et al., 2021). However, the unit of analysis, sample selection, empirical strategy, and ability to make causal claims have differed between studies. Scholars have used cities, census tracts, census blocks, HOLC areas, property addresses, and buffer zones as the unit of analysis. They have also focused on a single or several cities and have relied on observational and quasi-experimental designs.

Studies conducted at the city level leverage HOLC's non-random decision of creating appraisal maps only for cities above 40,000 residents (Anders, 2023; Faber, 2020). They use a regression discontinuity

[2] To be clear on the role of these institutions, HOLC focused on refinancing home mortgages, while the FHA concentrated on providing new home mortgages.

design by comparing cities just slightly above and below this population threshold. These studies provide credible causal claims on the effect of having residential security maps. Still, they estimate the net city-wide impact rather than the neighborhood effect. This approach implies that they cannot distinguish between new crimes due to concentrated community disadvantage and within-city redistribution of crimes.

Studies using the entire HOLC color-grade areas (Hoffman et al., 2020) compare large geographical units. Observable and unobservable confounders could affect the relationship between the color-grade categories and the outcome of interest. Other studies use census tracts as their unit of analysis in a random effects model framework (Faber, 2021; Hillier, 2003; Jacoby et al., 2018; Lyons et al., 2023; Mitchell and Chihaya, 2022). These studies measure the specific color-grade effects of the appraisal maps, but they have three main drawbacks. First, there is no one-to-one relationship between census tracts and color-grade areas. Researchers link the tracts by their geometric centroid, population-weighted centroid, or overlapping area, but they risk introducing non-random measurement errors in the variables, biasing the estimates. The linking process is relevant as there seems to be an optimal linking classification from a prediction perspective (Noelke et al., 2022), but its optimality for causal inference is unclear. Second, when comparing large areas with distinct observable characteristics (e.gr., race, ethnicity, and housing conditions), there are likely unobservable factors (e.gr., residents' preferences, risk attitudes, and networks) driving the differential outcomes, biasing the estimates. Finally, random effects models assume that the unobserved heterogeneity and the outcome of interest are uncorrelated. This strong assumption likely does not hold, which is the main reason for using quasi-experimental methods to estimate causal impacts.

The previous concerns also apply to the research using census block units (Appel and Nickerson, 2016; Benns et al., 2020; Poulson et al., 2021). Although these potential issues are less influential using smaller geographical areas, the empirical strategy matters. Studies using random effects (Benns et al., 2020; Poulson et al., 2021) unlikely can make causal claims. In contrast, quasi-experimental designs (Appel and Nickerson, 2016) comparing a discontinuous spatial change in the color-grade assignment at the border have a high interval validity to causal evidence to the extent that the identification assumption holds.

A strand of the literature uses small buffer zones near the boundaries to reduce the concerns of unmeasured confounders introducing endogeneity biases. These studies rely on a spatial regression discontinuity (Anders, 2023) or a difference-in-differences design compounded with a propensity score matching (Aaronson et al., 2021a,b). This last method provides insights into how the impacts change over time, but it requires highly granular data from over a century, which is rarely available, particularly for crime research.

Studies on crime and violence tend to find a worsening of public safety conditions in lower-grade categories, which is consistent with the negative effect caused by neighborhood disadvantage. However, research analyzing firearm violence (Benns et al., 2020; Jacoby et al., 2018; Poulson et al., 2021) and fatal police encounters (Mitchell and Chihaya, 2022) compare census blocks or tracts in a random effects model finding that redlining correlates to fatal violent increases of 70% to 700%. Accounting for post-intervention controls, violent crime rises by 30% (Lyons et al., 2023). A quasi-experimental design study revealed a 67% increase in violent crime (Anders, 2023). The large magnitude of the correlational studies is unsurprising as they do not overcome the negative self-selection bias, which leads to overestimating the impacts on crime.

In summary, the literature finds agreement on the long-lasting harmful effects of institutionalized racial discrimination caused by residential security maps. These results are consistent with the intergenerational effects research arguing that your place of origin matters on your long-term success (Chetty et al., 2016; Chetty and Hendren, 2018; Sharkey, 2008). Accordingly, there are reasons to believe law enforcement could also be affected.

3 Data

3.1 Data sources

This research hand-collected incident-level public safety information from 48 of the 203 US cities with a digitalized residential security map.[3] **Figure 1** shows the geographical distribution of these cities. Their spatial dispersion follows the US population density patterns.

The public safety administrative records come from each city's police department. The crimes and arrests are categorized into index, low-level, and weapon-related offenses. Violent (murder, robbery, and aggravated assault) and property (burglary, theft, and motor vehicle theft) offenses represent the index category. Low-level offenses are all the other incidents reported to law enforcement. Dangerous weapon offenses include gun discharges, illegal criminal possession, shootings, and brandishing. This research also collects information on police stops and 911 calls for service. Arrests and police stops are a measurement of police proactivity. Emergency calls for service reflect basic citizens' cooperation with police and perceptions of legitimacy and trust, as without it, criminal activity would be severely underreported.

Three cities do not disclose the location of homicides, and one of them of aggravated assaults. All other

[3]Not all cities have a digitalized residential security map. The 48 cities included in the study are Asheville, NC; Atlanta, GA; Austin, TX; Baltimore, MD; Birmingham, AL; Boston, MA; Buffalo, NY; Cambridge, MA; Charlotte, NC; Chattanooga, TN; Chicago, IL; Columbus, OH; Dallas, TX; Denver, CO; Detroit, MI; Durham, NC; Fort Worth, TX; Greensboro, NC; Hartford, CT; Houston, TX; Indianapolis, IN; Kansas City, MO; Lincoln, NE; Little Rock, AR; Los Angeles, CA; Louisville, KY; Memphis, TN; Milwaukee, WI; Minneapolis, MN; Montgomery, AL; Nashville, TN; New Orleans, LA; New York, NY; Norfolk, VA; Oakland, CA; Omaha, NE; Philadelphia, PA; Phoenix, AZ; Pittsburgh, PA; Portland, OR; Rochester, NY; Saint Louis, MO; Saint Paul, MN; San Francisco, CA; Seattle, WA; Somerville, MA; Syracuse, NY; Tulsa, OK.

cities report all index crimes. Low-level offenses and dangerous weapon crimes are reported by 42 and 38 cities. Arrest and police stop data are less frequently available as only 23 and 25 cities release incident-level georeferenced information. Calls for service are reported by 18 cities. **Appendix Figure A.1** identifies the data availability by city and public safety outcome.

The digitalized residential security maps were retrieved from Mapping Inequality (Nelson et al., 2023). Only areas within the local police department jurisdiction (e.g., city limits) were included in the analysis. **Appendix Figure A.2** shows the residential security maps used in this research. A visual inspection reveals that the color-grade categories tend to be clustered within cities. The individual borders of each category tend to be idiosyncratically defined (e.gr., they have very irregular shapes in some cities, while others use rectangular ones). The color areas were drawn only in residential neighborhoods. Sparsely built-up developments, farmland, industrial, and commercial areas were not assigned a category in the 1930s.

The historical decennial census was retrieved from the Integrated Public Use Microdata Series USA data (Ruggles et al., 2021). Respondent-level data with available street address information is only available to a sample of the decennial census. Mainly, it was used the 1900 1.2%, 1910 1.4%, 1920 1%, and 1930 5% decennial census samples.[4]

3.2 Analytical database

The analytical database uses public safety data from 2014 to 2019. This period was chosen as most cities have information from 2014 to the present. It also excludes any post-COVID-19 periods to avoid analyzing atypical public safety patterns. More than half of the 48 cities included in this research only release the address of the public safety incident, usually rounding the last address digits to the nearest hundred block to protect victims' privacy. After performing basic data cleaning processes and using alternative geocoders to increase the probability of obtaining the latitude-longitude values, a manual random inspection of the incidents revealed that the geocoding hit rate was above the minimum acceptable threshold (Ratcliffe, 2004). Likewise, the crimes were compared to the Federal Bureau of Investigation's Uniform Crime Reporting system to ensure accuracy. The crime data matched well in levels and trends. Historical decennial census data was also geocoded using the available address information.

The unit of analysis could affect the presence of confounders biasing the relationship. The spatial data

[4] The decennial census data is a sample, so there are missing values across several hexagons. To avoid estimating changes due mostly to sample size differences, whenever historical census data was used as a control variable, the missing data was extrapolated using the mean value of its neighboring hexagons. The results are qualitatively similar using the raw census data. Models using the census data as the dependent variable use the raw data to avoid introducing measurement errors.

aggregation process can also impose challenges due to the modifiable areal unit problem noted by the crime and cartographic literature (Bernasco and Elffers, 2010; Openshaw, 1984; Ratcliffe, 2010). To reduce such concerns, a hexagonal grid of 100 meters (328 feet) per side (25,980 m^2 or 279,646 square feet) was created over the color-grade areas to ensure homogeneous spatial units. The hexagons' centroids were overlapped with the HOLC maps to identify their location. They were also used to estimate the Euclidean distance to the closest border of a different HOLC grade. The small size of the hexagons drastically reduces the concern of overlapping multiple color-grade categories. Although it may introduce a classical random measurement error, reducing the precision of the estimates but not biasing them (Wooldridge, 2010), particularly when comparing very small distances to the border. The present-day (2014-2019) public safety and historical (1900-1930) decennial census data were pooled and aggregated to the hexagonal level, creating a cross-sectional dataset.

Table 1 exhibits descriptive statistics of the areas within 500 meters of the border of a color-grade category. Historical data shows that the areas that eventually would experience a color-grade category had sociodemographic differences. White residents are the most common racial group in all areas, with nearly 96 percent of the population, but their proportion is smaller in red areas (87 percent). The Black population is four times larger in the red category than the other categories: it accounts for 12.2 percent of the population in the average red hexagon between 1900 and 1930. The difference between Asian and Native American individuals across categories is small (0.01 to 0.03 percent). There are almost twice the number of foreign-born residents in the red area compared to the green area (12.0 vs 19.6 percent). Homeownership rates are different between the green and red categories (65 vs 35.3 percent). The number of people who cannot read and write also is four times larger between the green and red areas (0.7 vs 3.0 percent).

Present-day public safety data worsens as the grade moves down the classification ladder (green to red categories). The average hexagon in a green area within 500 meters its border experienced 9.1 index crimes between 2014 and 2019, while its counterpart red grade had twice the criminal offenses (22.3 offenses). Thefts are the most commonly reported offense, followed by burglaries and aggravated assaults. Murders, weapons violations, and robberies are infrequent crimes. Low-level offenses are four times larger between the green and red categories (10 and 43.6 low-level offenses), while dangerous weapons offenses are more than 10 times larger (0.1 and 1.1 offenses). These patterns are consistent with national crime data.

Law enforcement agencies are more likely to clear violent than property crimes by arresting a suspect. The most common arrest categories among index crimes are aggravated assaults, thefts, and robberies. There are more arrests in the red category than in the other color-grade hexagons. For instance, there were around 0.2 aggravated assaults in the average green hexagon, while this figure is 2.1 in the red category. Low-level offense arrests also experience a similar trend: there are 31.3 in the red category, while only 6.8 arrests are in the green one. Another indicator measuring law enforcement activity is police stops. A typical hexagon experienced 15.3 police stops in a green area between 2014 and 2019. This number is 61.5 in the red category. Finally, calls for service are three times in the red relative to the green category (114 and 378). In summary, the descriptive statistics show that the red areas experience more serious and non-serious crimes and police activity.

4 Empirical strategy

The previous section shows that the areas defined as "obsolete" and "undesirable" in the 1930s based on the discriminatory redlining criteria are more unsafe nowadays. Did redlining worsen their public safety conditions? A major concern in answering this question is that pre-existing unaccounted factors could have caused the public safety differences, even if the residential security maps had not been drawn. Such factors are plausible as the color classification was not random. Appraisers explicitly used race, ethnicity, and housing conditions to draw the maps, and they successfully achieved it as these places differed in such dimensions in the 1930s (Hillier, 2005). Likely, these areas had different propensities toward crime, so redlining could have made explicit pre-existing public safety differences across places rather than causing them. Said differently, unsafe areas in the 1930s could still be unsafe nowadays, independently of the residential security maps' existence, as crime concentration is relatively stable across time (Curman et al., 2015; Weisburd et al., 2004).

To address these endogeneity concerns, this research leverages the spatial discontinuous change in lending decisions caused by the redlining maps following previous studies in this area (Anders, 2023; Appel and Nickerson, 2016; Fishback et al., 2020). The research design finds motivation from uncertainty around where to place the color-grade boundaries. HOLC elicited participation from local stakeholders and consultants to create the maps. They may have agreed on classifying large areas (e.g., identifying the south of any city as high-minority prevalent or the west side as poorly built). However, it is possible they were uncertain about where *exactly* to draw the borders. Some appraisers may have thought it should be south of 11th street, others south the 15th or 4th street of any given city. This uncertainty comes from different beliefs, expectations, and perceptions about the areas. This situation is credible as the classification did not rely on existing administrative or political boundaries.

Moreover, the maps were done in several iterations, highlighting that appraisers disagreed on defining the boundaries (Hillier, 2005). Nevertheless, once the maps were defined, they may have aligned the stakeholders' beliefs about which neighborhoods were "on-demand" (green), "still good" (blue), "almost obsolete" (yellow), and "undesirable" (red), causing a sharp spatial discontinuity in lending decisions for everyone where it did not use to exist. Consequently, by comparing areas near the boundaries of different color-grade categories, this research measures the long-term effects of redlining on public safety.

The econometric specification uses the following spatial regression discontinuity design, restricting the sample within a small bandwidth h:

$$y_i = \alpha_0 + \beta_1 D_i + \alpha_1 f(r_i) + \alpha_2 D_i g(r_i) + u_i \tag{1}$$

where y_i is the count of public safety incidents (e.gr., crimes, arrests, police stops, or calls for service) between 2014 and 2019 in hexagon i. D_i is an indicator variable for being in a low-grade area relative to its neighboring border: a red grade adjacent to a yellow, a yellow grade adjacent to a blue, and a blue grade adjacent to a green. The running variable, r_i, is the distance to the closest border of a different color-grade category, centered around zero, so a negative distance means being in the higher category and a positive in the lower grade. The functions $f(r_i)$ and $g(r_i)$ allow flexible polynomial forms on either side of the border. The error term is represented by u_i, using robust standard errors. β_1 is the main coefficient of interest and provides the average effect for areas that would have been classified as a higher category had they been on the other side of the threshold. Given the non-negative nature of the count of public safety incidents and the common presence of zeros in small geographical areas, particularly for serious offenses, equation 1 uses a Poisson regression model for public safety outcomes. The estimate is interpreted as proportional changes. When the outcome variable is a sociodemographic variable measured in percent terms, the model uses an Ordinary Least Squares regression to measure changes in percentage points. Using hexagons to aggregate uneven spatial evens contributes to ensuring the running variable distribution is smooth around the threshold.[5]

The spatial regression discontinuity design resembles the method used to identify the long-term treatment effects of institutionalizing systems based on geographical locations (Dell, 2010; Shertzer et al., 2018). It has been used to estimate the crime treatment effects of legalizing marijuana (Dragone et al., 2019), urban development (Mitre-Becerril and MacDonald, 2024), driving restrictions (Carrillo et al., 2018), and private policing (MacDonald et al., 2016).

[5] **Appendix Figure C.1** shows that the density of tracts near the threshold is similar on either side.

5 Results

This section presents the estimates of redlining on pre-intervention historical sociodemographic outcomes, followed by present-day public safety and law enforcement measured by criminal offenses, arrests, police stops, and calls for service.

5.1 Historical outcomes

This research first provides descriptive evidence of differences in variables measuring racial, ownership, and disadvantaged disparities that the real estate agents could have used to develop the HOLC maps. **Figure 2** compares mean pooled hexagonal decennial data between 1900 and 1930 (pre-intervention period) using bins of 15 meters in a 700-meter bandwidth (around seven city blocks) near the border of different color-grade categories. Second-order polynomials fit the data points. A negative distance means being in the higher category and a positive in the lower grade. There are small but discrete changes at the border. The proportion of Black, foreign-born, and illiterate population increased, while homeowners decreased in areas that would eventually experience a lower category relative to their adjacent counterparts of a higher color grade.

These results are confirmed using the spatial regression discontinuity estimated via ordinary least squares with a second-order polynomial function, robust-standard errors, and the optimal bandwidths based on Imbens and Kalyanaraman (2012) and Calonico et al. (2015). **Table 2** shows a statistically significant one to two percentage points increase in the proportion of Black residents in areas that would become lower grade categories relative to their higher grade neighbors. This effect translates into a 15 to 32 percent change depending on the chosen bandwidth (the mean outcome baseline is six percent). Residents also have a half percentage point increase in the probability of being illiterate (can't read nor write), which translates into a 26 percent increase. The effect in foreign born is also positive but is not statistically significant due to its large standard errors. Residents are three percentage points or 7.2 percent less likely to be homeowners in areas that eventually would become low grade categories (the baseline probability is 44 percentage points). The effect on illiterate and homeownership are only statistically significant using a larger bandwidth, probably due to statistical power issues. While the coefficient of foreign-born residents is positive, the large standard errors limit making any statements about its impact.

These results suggest that stakeholders, map consultants, and appraisals strategically chose areas to impact home mortgage lending. These results are consistent with recent research (Fishback et al., 2023), casting doubts on the quasi-random assignment of delineating the color-grade borders. The authors also found differential effects in neighboring areas before the maps were developed when comparing data from 10 cities.

5.1.1 Present-day reported public safety

This section also presents first descriptive evidence on whether public safety drastically changes at the border by comparing the mean hexagonal crimes between the lower and higher color categories. **Figure 3** shows the scatter plot and second-order polynomials fit of the count of crimes in 15-meter bins. There is a discontinuity at the border for all crime categories: violent, property, low-level, and dangerous weapons offenses increase in the lower category relative to their adjacent counterparts (e.gr., a red grade relative to yellow, a yellow grade relative to blue, and a blue grade relative to green).

Table 3, Panel A confirms these results using a spatial regression discontinuity with a second-order polynomial function, robust-standard errors, and the optimal bandwidth following Calonico et al. (2015). Index offenses are 12.3 percent larger (2.3 more criminal offenses per hexagon between 2014 and 2019) in lower color areas relative to their adjacent higher-grade counterparts. One hexagon is comprised of around one city block. This increment comes from having more violent and property crimes, which increased by 14.9 and 11.2 percent (0.6 and 1.6 more offenses). The crime subcategories also experienced significant increases in the average hexagon: murder (32% or 0.03 offenses), robbery (17.7% or 0.3 offenses), aggravated assault (10.1% or 0.3 offenses), burglary (10.1% or 0.3 offenses), theft (11% or 0.9 offenses), and motor vehicle theft (11.7% or 0.3 offenses). There was also an increase in low-level offenses (12.6% or 3.9 offenses) and dangerous weapons (18.1% or 0.1 offenses).

These previous estimates would have been accurate estimations of the effects of redlining were there no differences in pre-intervention outcomes at the border between lower and higher grade areas. However, there are discrete changes in pre-intervention sociodemographics. This situation raises limitations for the spatial regression discontinuity design as one identifying assumption of the model is the lack of other changes at the same cutoff value (Imbens and Lemieux, 2008). To account for this concern, **Table 3**, Panel B estimates the regression discontinuity by adding the historical pre-HOLC decennial census controls to account for differential sociodemographic compositions around the threshold. The coefficients on the effects of downgrading an area shrink by one-third, on average (outcomes shrink between one-fifth and half). Still, all the estimates remain statistically significant.

By adding pre-intervention controls, index crimes now show an increase of 7.5% (1.7 offenses) caused by a 9.3% (0.5 offenses) and 7.0% (1.2 offenses) increase in violent and property crimes. All violent crime subcategories also increased in smaller magnitudes: murder (23.3% or 0.02 offenses), robbery (10.6% or

0.2 offenses), and aggravated assault (7.9% or 0.25 offenses). Property crimes experience the same trend: burglaries (8.3% or 0.3 offenses), larceny (6% or 0.7 offenses), and motor vehicle thefts (8% or 0.2 offenses) show significant increases between 2014 and 2019 for being in lower grade categories relative to their neighboring counterparts. Likewise, low-level and dangerous weapons violations now show a 7.1% (2.9 offenses) and 14.7% (0.15 offenses) increment.

Table 4 and **Appendix Figure B.1** shows a spatial regression discontinuity of the effects of redlining on law enforcement measured through police arrests. All coefficients have positive signs for the model without census controls but are mostly not statistically significant. Several coefficients flip sign and magnitude once controlling for pre-intervention covariates. The only exception is arrests for dangerous weapons violations. It shows a significant increase of 27.7% that decreases to 22.1% once controlling for historical census variables, which translates to an additional 0.24 to 0.26 offenses per hexagon between 2014 and 2019. Dangerous weapons arrests are commonly a by-product of proactive enforcement (Braga et al., 2022; Moore, 1980), suggesting that police activity is influenced by redlining.

Another way to measure police proactivity and citizens' cooperation with law enforcement is by using police stops and calls for service. **Appendix Figure B.2** shows suggestive evidence that there are more police stops in areas that were categorized as a lower grade. **Table 5** presents the spatial regression discontinuity coefficient. There is a 16.2% increase (8.5 incidents) in police stops due to redlining after controlling for historical census sociodemographics. The results are imprecisely estimated, and it is significant without the census controls at the 0.109 alpha value. Note that the point estimate is the same across specifications, but the standard errors change. Calls for service show a marginally significant decrease of 9.9% or around 331 fewer 911 calls done in the lower grade categories relative to their higher grade neighbors. The result is not significant without historical census controls.

In summary, present-day serious and non-serious criminal activity increased due to redlining. However, without controlling for initial differences in the areas surrounding the color-grade categories, the effects are overestimated. This result is consistent with previous studies (using random effects and census tracts) finding that once controlling for historical census data, the estimates become smaller (Jacoby et al., 2018; Lyons et al., 2023). Law enforcement proactivity measured through arrests of serious and non-serious offenses are mostly null effects. Only arrests for dangerous weapons violations, a common result of proactive enforcement, show a consistently significant increase. The impacts on police stops and calls for service, though imprecisely measured, show that redlining also affected them in the long run.

5.2 Robustness

To ensure the main results are not an artifact of random chance, the results are examined under alternative specifications and placebo tests. The econometric model relaxes the assumptions by changing the polynomial function, the bandwidth around the threshold, the use of sociodemographic controls, and, for this particular research, city-fixed effects. Given the many combinations of the different assumptions, this research uses a specification chart. This figure helps to identify whether any consequential analytical decisions could influence the findings and provides a more precise statistical inference of the econometric modeling (Simonsohn et al., 2020). **Appendix Figure C.2** displays the specification chart by including the distribution of estimates and confidence intervals under different assumptions.

First, the regression model changes the functional form for the running variable, $f(r_i)$ and $g(r_i)$ in equation 1, using a linear, quadratic, and cubic polynomial function. Using higher polynomial functions slightly increases the standard errors, but the point estimates remain relatively the same. For example, index crimes have point estimates between 0.048 and 0.053 (5.0% to 5.4% change) and standard errors between 0.016 and 0.028 (1.7% to 2.9%). The lower precision still keeps the statistical significance below a 0.1 pvalue. Even murder, the rarest but most costly criminal offense, shows no meaningful changes under different polynomial functions. The coefficient goes from 0.147 to 0.164 (15.9% to 17.8%), and its standard errors range between 0.039 and 0.070 (4.6% to 8.3%). The remaining crime outcomes face the same trend using alternative polynomial functions.

A different analytical choice is the bandwidth selection. The model is estimated under different bandwidths, 500 to 2,000 meters, in 250 increments. This distance equals four to sixteen blocks away from the downgraded border. The results follow the bias-variance trade-off where smaller bandwidths are more likely to remove unobserved endogeneity biases at the cost of increasing the imprecision of the estimates. For example, the estimates for property crime less than 900 meters away from the border are, on average, 0.040 (4.0% change), increasing to 0.046 (4.7%) when using 1,700 to 2,000-meter bandwidths. In contrast, the standard errors decrease from 0.028 (3.0%) to 0.019 (2.0%). A similar situation happens with all other crimes.

Third, it has been shown that using historical decennial census controls in the specification matters. Adding controls shrinks the estimates. This result also holds when changing simultaneously the bandwidth and polynomial functions. Adding historical census data does not shrink the standard errors, which explains why some estimates now become significant at the 0.1 value.

Next, the data was compiled from 48 police departments. A regression discontinuity does not require fixed effects for identification. Still, one could be concerned that the model compares areas from different cities. To ensure the comparison is done within cities, an alternative specification adds city-fixed effects. The points estimates become slightly smaller, but also the standard errors. For example, the dangerous weapons offense coefficient goes from 0.141 to 0.135 (15.1% to 14.5%), and the standard errors also decrease from 0.061 to 0.059 (7.0% to 6.8%). This result is unsurprising as fixed effects capture any observed variation that may explain the outcome, along with sample variance.

The coefficient sign flips in 0.6% of the 935 alternative specifications ran across the eleven crime categories. It means that only very specific analytical choices could imply that crime did not increase due to redlining, although none are statistically significant. Accordingly, the main takeaway from the crime specification chart is that the point estimate and statistical significance may change slightly across analytical choices. Still, the results suggest that redlining is correlated to an increase in crime even after seven decades of its implementation and repeal.

Moving on to the arrests robustness checks, **Appendix Figure C.3** show the estimates under similar alternative specifications. Most specifications are around zero with large confidence intervals, and signs flip, limiting the ability to detect changes due to redlining. The only outcome that consistently has a positive effect is arrest for dangerous weapons. Across all the different model specifications, 59 percent of weapon arrest coefficients are significant at a 0.05 alpha level, while this number increases to 86 percent for a 0.1 alpha level. These results suggest that redlining impacts police proactivity measured through an arrest for weapons violations.

Next, **Appendix Figure C.4** shows the specification chart for police stops and calls for service. These outcomes had a marginally significant effect on their preferred specification. The robustness checks suggest that calls for service are not robust to different model choices, as its sign flips in 12% of the 85 alternative models. None of the estimates are significant at a 0.05 alpha level, and only 2.3 percent are significant at a 0.1 alpha level. In contrast, police stops are more robust. All its estimates are positive, and 24.7% of the alternative specifications are significant at a 0.05 alpha level, and this number increases to 45.9% for a 0.1 alpha level.

Finally, one concern of the identification strategy is the existence of discontinuities at other values besides the border. Following the regression discontinuity literature (Imbens and Lemieux, 2008), **Appendix Figure C.5** test for jumps at non-discontinuity values by shifting the threshold of the running variable. It assumes the border is located from a negative to a positive 2,100 meters relative to the actual border in 700 increments. Then, it estimates the regression at the optimal bandwidth at the new threshold for the following outcomes: the eleven crime categories, police stops, and arrests for weapon violations (the outcomes with significant effects). Thirteen outcomes and six alternative new threshold values result in 78 estimates. Out of these 78 estimates, 12 coefficients are statistically significant. While this rate is slightly larger than a conventional false discovery rate, it is common in a multi-testing hypothesis (e.gr., an increase in robbery influences violent crime, which also affects index crimes, so the outcomes are not statistically independent). Equally importantly, the estimates are centered around zero, suggesting that the effects only happen at the real border.

5.3 Heterogeneity

Previous evidence highlights that redlining affected communities by reducing homeownership, house values, rents, and income, while also increasing racial segregation and poverty across time (Aaronson et al., 2021a,b; Faber, 2020). These factors can influence public safety. The lack of mortgages and homeownership can increase crime (Bunting, 2020; Disney et al., 2023; Mitre Becerril, 2023). Property prices and crime are negatively correlated (Pope and Pope, 2012), while the evidence on income, poverty, and public safety is longstanding (Sharkey et al., 2016).

This research lacks pre-intervention data to disentangle which factors are worsening neighborhood public safety. Nevertheless, it can examine heterogeneous responses across places. Specifically, it assesses whether pre-intervention baseline differences are correlated with larger crime impacts. To achieve this goal, all the parameters of the regression discontinuity are interacted with a relevant heterogeneity dimension from the pre-HOLC maps period (1900-1930): i) being in a hexagon where more than half of the residents are homeowners, 2) being in a hexagon where more than half of the residents are foreign-born, and 3) being in a hexagon where there is a strictly positive number of non-white residents. The econometric specification restricts the sample within a small bandwidth h and is as follows:

$$y_i = \alpha_0 + \alpha_1 f(r_i) + D_i(\alpha_2 + \alpha_3 f(r_i)) + H_i[\alpha_4 + \alpha_5 f(r_i) + D_i(\beta_1 + \alpha_6 f(r_i))] + e_i \tag{2}$$

where H_i is the pre-HOLC map heterogeneity variable measured as an indicator variable, and β_1 is the coefficient of interest. The rest of the parameters are as explained in the previous section. This specification leverages high versus low baseline demographic differences and discontinuous spatial variation of the redlining maps. This specification resembles a differences-in-discontinuities.[6] Intuitively, we can think of this model as examining whether the main coefficient of two regression discontinuities is different

[6] The differences-in-discontinuities design usually uses pre-post changes (as in a difference-in-differences) along a border comparison (as in a regression discontinuity). Still, the econometric specification is the same. The model has been used in the economics and crime literature (Grembi et al., 2016; Mitre-Becerril and MacDonald, 2024).

between them (but everything in one single equation).

Table 6 displays β_1 from Equation 2. Panel A, the model using homeowner as the heterogeneity dimension, shows statistically significant impacts for violent crime and its subcategories (murder, robbery, and aggravated assault). The effect on property crime is positive but not statistically significant, except for motor vehicle theft. This result suggests that among places affected by redlining, those with high homeownership rates experienced larger crime increases than those with fewer homeowners. It may be that these places had different characteristics that made them adjust differently to redlining. This project is limited to provide more granular insights but it is an area of future research.

Panel B uses foreign-born as the heterogeneity dimension. Places with a higher proportion of residents born abroad in 1990-1930 seem to have experienced larger increases in burglary and motor vehicle theft crimes than places with fewer foreign-born individuals. The coefficients for violent crime are positive, but none are statistically significant. It may be that the country of origin matters for how residents adapted to the neighborhood disinvestment. Finally, Panel C shows no significant changes when using non-white as the heterogeneity dimension.

6 Discussion and conclusion

This research contributes to the literature on the long-term effects of structural discrimination on public safety. It examines changes in criminal activity and law enforcement caused by the residential security maps enacted almost a hundred years ago. It leverages within-city variation in a spatial regression discontinuity framework by comparing areas near different color-grade boundaries within a small distance (around six to ten city blocks). Decennial census data from 1900 to 1930 on the proportion of Black, homeowners, and illiterate populations, reveal discrete changes at the border before the maps were implemented. The pre-intervention differences influence the regression discontinuity estimates for present-day outcomes. To address these concerns, the econometric specifications control for those historical sociodemographic differences.

Using incident-level data from 48 US cities, the regression discontinuity shows significant increases between six to 15 percent in violent, property, low-level, and dangerous weapon offenses in areas labeled as a lower grade relative to their similarly situated counterparts with a higher category once accounting for the historical differences. Had the specification not controlled for the pre-intervention sociodemographics, the estimates would have been sixty percent larger. The preferred estimates are considerably smaller than previous studies have suggested. For instance, firearm violence has been estimated to be 70 to 700 percent larger due to redlining (Benns et al., 2020; Jacoby et al., 2018; Poulson et al., 2021), while this research suggests a 15 percent change. Even recent evidence (Lyons et al., 2023) assessing violent crime with a fully saturated model shows a 30 percent change, contrasting with the 9.3 percent estimated by this research. Overall, the results suggest that redlining is an important part of historical and persistent neighborhood disinvestment and inequality in the US, but its impact on present-day crime may have been overestimated by previous studies.

One of the contributions of this research is the estimation of redlining on police activity. There is a 22 percent increase in arrests for dangerous weapon violations. This type of arrest is a common by-product of police stops. No other arrest category (violent, property, or low-level offenses) shows a consistent significant change. Police stops increased by 16 percent due to redlining, although it is imprecisely measured. Emergency 911 calls for service seem to have decreased due to redlining, but the estimate is not robust to all the alternative specification checks. These results suggest that police proactivity may have been impacted by redlining. Previous studies have suggested that racial segregation increased due to redlining (Faber, 2020). Accordingly, it may be that law enforcement may be behaving differently due to the neighborhood composition. Future research should explore the mechanism behind these results.

This research faces is constrained to provide evidence on the causal mechanisms. Nevertheless, it highlights heterogeneous dosage responses across places with different baseline sociodemographic compositions. The results highlight that places with higher homeownership rates experienced larger violent crime increases due to redlining than those with fewer homeowners. In contrast, there is evidence that some property crime categories (burglary and motor vehicle theft) had larger impacts associated with redlining in areas with more foreign-born residents.

It is important to consider the limitations of this research. First, the 48 cities included in the analysis are the most populated cities with available public safety information and a residential security map. While they show a geographical distribution consistent with the US population density, they do not unequivocally represent the 239 cities with HOLC maps. Whether the effect may be different in the excluded cities is an empirical question that future research could address. Second, this research cannot provide the differential temporal evolution of crime across color-grade categories. Unfortunately, there is no georeferenced public safety repository from the last century. Having such data would provide insights into how long neighborhoods start worsening their public safety conditions due to community disadvantage. Future research could aim to collect crime data from historical archives. Third, the decennial census data are samples of the censuses from 1900 to 1930. Access to the full census would increase the sample size to detect smaller effects. Fourth, the crime data comes from 48 cities, but the arrest data only comes from 23 cities. These differences in sample size could limit the statistical power to detect small changes in police enforcement. Collecting more incident-level law enforcement data can help examine the extent to which police activity was affected by redlining.

Finally, this research highlights that the 1930s residential security maps exacerbated disadvantages, preventing safer communities in the long run. Nevertheless, redlining is one of the different urban economic events that have caused inequality and disadvantage to racial-minority groups. It is important that future studies also focus on how redlining interacted with more recent sources of racial inequality, and examine the cascading consequences of structural discrimination of place-based policies and economic restructuring on historically underrepresented communities.

References

Aaronson, D., Faber, J., Hartley, D., Mazumder, B., and Sharkey, P. (2021a). The long-run effects of the 1930s holc "redlining" maps on place-based measures of economic opportunity and socioeconomic success. *Regional Science and Urban Economics*, 86:103622.

Aaronson, D., Hartley, D., and Mazumder, B. (2021b). The effects of the 1930s holc" redlining" maps. *American Economic Journal: Economic Policy*, 13(4):355–92.

Alpert, G. P., MacDonald, J. M., and Dunham, R. G. (2005). Police suspicion and discretionary decision making during citizen stops. *Criminology*, 43(2):407–434.

Anders, J. (2023). The long run effects of de jure discrimination in the credit market: How redlining increased crime. *Journal of Public Economics*, 222:104857.

Appel, I. and Nickerson, J. (2016). Pockets of poverty: The long-term effects of redlining. *Available at SSRN 2852856*.

Bartik, T. J. (2020). Using place-based jobs policies to help distressed communities. *Journal of Economic Perspectives*, 34(3):99–127.

Beck, B. (2020). Policing gentrification: Stops and low–level arrests during demographic change and real estate reinvestment. *City & Community*, 19(1):245–272.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of political economy*, 76(2):169–217.

Benns, M., Ruther, M., Nash, N., Bozeman, M., Harbrecht, B., and Miller, K. (2020). The impact of historical racism on modern gun violence: redlining in the city of louisville, ky. *Injury*, 51(10):2192–2198.

Bernasco, W. and Elffers, H. (2010). Statistical analysis of spatial crime data. *Handbook of quantitative criminology*, pages 699–724.

Braga, A. A., Brunson, R. K., and Drakulich, K. M. (2019). Race, place, and effective policing. *Annual review of sociology*, 45:535–555.

Braga, A. A., Cook, P. J., and Douglas, S. (2022). Prevention strategies for policing gun violence. *The ANNALS of the American Academy of Political and Social Science*, 704(1):158–180.

Brantingham, P. J. and Uchida, C. D. (2021). Public cooperation and the police: Do calls-for-service increase after homicides? *Journal of criminal justice*, 73:101785.

Brunson, R. K. and Weitzer, R. (2009). Police relations with black and white youths in different urban neighborhoods. *Urban Affairs Review*, 44(6):858–885.

Bunting, W. (2020). Does increased access to home mortgage money reduce local crime rates? evidence from san diego county. *Regional Science and Urban Economics*, 84:103570.

Bursik Jr, R. J. (1988). Social disorganization and theories of crime and delinquency: Problems and prospects. *Criminology*, 26(4):519–552.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2015). Rdrobust: an r package for robust nonparametric inference in regression-discontinuity designs. *R J.*, 7(1):38.

Carrillo, P. E., Lopez-Luzuriaga, A., and Malik, A. S. (2018). Pollution or crime: The effect of driving restrictions on criminal activity. *Journal of Public Economics*, 164:50–69.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, 18(1):234–261.

Chalfin, A., Hansen, B., Weisburst, E. K., and Williams Jr, M. C. (2022). Police force size and civilian race. *American Economic Review: Insights*, 4(2):139–58.

Chetty, R. and Hendren, N. (2018). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Clarke, R. V. and Cornish, D. B. (1985). Modeling offenders' decisions: A framework for research and policy. *Crime and justice*, 6:147–185.

Curman, A. S., Andresen, M. A., and Brantingham, P. J. (2015). Crime and place: A longitudinal examination of street segment patterns in vancouver, bc. *Journal of Quantitative Criminology*, 31:127–147.

Dell, M. (2010). The persistent effects of peru's mining mita. *Econometrica*, 78(6):1863–1903.

Disney, R., Gathergood, J., Machin, S., and Sandi, M. (2023). Does homeownership reduce crime? a radical housing reform from the uk. *The Economic Journal*, 133(655):2640–2675.

Dragone, D., Prarolo, G., Vanin, P., and Zanella, G. (2019). Crime and the legalization of recreational marijuana. *Journal of economic behavior & organization*, 159:488–501.

Faber, J. (2021). Contemporary echoes of segregationist policy: Spatial marking and the persistence of inequality. *Urban Studies*, 58(5):1067–1086.

Faber, J. W. (2020). We built this: Consequences of new deal era intervention in america's racial geography. *American Sociological Review*, 85(5):739–775.

FHA (1936). Underwriting manual: underwriting analysis under title ii. Technical report, Federal Housing Administration.

Fishback, P. V., Flores-Lagunes, A., Horrace, W. C., Kantor, S., and Treber, J. (2011). The influence of the home owners' loan corporation on housing markets during the 1930s. *The Review of Financial Studies*, 24(6):1782–1813.

Fishback, P. V., LaVoice, J., Shertzer, A., and Walsh, R. (2020). The holc maps: How race and poverty influenced real estate professionals' evaluation of lending risk in the 1930s. Technical report, National Bureau of Economic Research.

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

Grembi, V., Nannicini, T., and Troiano, U. (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics*, pages 1–30.

Hillier, A. E. (2003). Redlining and the home owners' loan corporation. *Journal of Urban History*, 29(4):394–420.

Hillier, A. E. (2005). Residential security maps and neighborhood appraisals: the home owners' loan corporation and the case of philadelphia. *Social Science History*, 29(2):207–233.

Hoffman, J. S., Shandas, V., and Pendleton, N. (2020). The effects of historical housing policies on resident exposure to intra-urban heat: a study of 108 us urban areas. *Climate*, 8(1):12.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.

Jackson, K. T. (1980). Race, ethnicity, and real estate appraisal: The home owners loan corporation and the federal housing administration. *Journal of Urban History*, 6(4):419–452.

Jacoby, S. F., Dong, B., Beard, J. H., Wiebe, D. J., and Morrison, C. N. (2018). The enduring impact of historical and structural racism on urban violence in philadelphia. *Social Science & Medicine*, 199:87–95.

Jones-Correa, M. (2000). The origins and diffusion of racial restrictive covenants. *Political Science Quarterly*, 115(4):541–568.

Kelling, G. L., Wilson, J. Q., et al. (1982). Broken windows. *Atlantic monthly*, 249(3):29–38.

Ladd, H. F. (1994). Spatially targeted economic development strategies: do they work? *Cityscape*, 1(1):193–218.

Laniyonu, A. (2018). Coffee shops and street stops: Policing practices in gentrifying neighborhoods. *Urban affairs review*, 54(5):898–930.

Lyons, C. J., Vélez, M. B., and Chen, X. (2023). Inheriting the grade: Holc "redlining" maps and contemporary neighborhood crime. *Socius*, 9:23780231231197030.

MacDonald, J. M., Klick, J., and Grunwald, B. (2016). The effect of private police on crime: evidence from a geographic regression discontinuity design. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 831–846.

Massey, D. S. (2015). The legacy of the 1968 fair housing act. In *Sociological Forum*, volume 30, pages 571–588. Wiley Online Library.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Merton, R. K. (1938). Social structure and anomie. *American sociological review*, 3(5):672–682.

Michney, T. M. and Winling, L. (2020). New perspectives on new deal housing policy: explicating and mapping holc loans to african americans. *Journal of Urban History*, 46(1):150–180.

Mitchell, J. and Chihaya, G. K. (2022). Tract level associations between historical residential redlining and contemporary fatal encounters with police. *Social Science & Medicine*, 302:114989.

Mitre Becerril, D. (2023). Banks against crime: The impact of home mortgages on neighborhood crime. Technical report, Working Paper.

Mitre-Becerril, D. and MacDonald, J. M. (2024). Does urban development influence crime? evidence from philadelphia's new zoning regulations. *Journal of Urban Economics*, 142:103667.

Moore, M. H. (1980). The police and weapons offenses. *The ANNALS of the American Academy of Political and Social Science*, 452(1):22–32.

Nardone, A., Casey, J. A., Morello-Frosch, R., Mujahid, M., Balmes, J. R., and Thakur, N. (2020). Associations between historical residential redlining and current age-adjusted rates of emergency department visits due to asthma across eight cities in california: an ecological study. *The Lancet Planetary Health*, 4(1):e24–e31.

Nelson, R. K., Winling, L., Marciano, R., and Connoly, N. (2023). Mapping inequality. Technical report, American Panorama].

Neumark, D. and Simpson, H. (2015). Place-based policies. In *Handbook of regional and urban economics*, volume 5, pages 1197–1287. Elsevier.

Noelke, C., Outrich, M., Baek, M., Reece, J., Osypuk, T. L., McArdle, N., Ressler, R. W., and Acevedo-Garcia, D. (2022). Connecting past to present: Examining different approaches to linking historical redlining to present day health inequities. *PloS one*, 17(5):e0267606.

Openshaw, S. (1984). The modifiable areal unit problem. *Concepts and techniques in modern geography.*

Pope, D. G. and Pope, J. C. (2012). Crime and property values: Evidence from the 1990s crime drop. *Regional science and urban economics*, 42(1-2):177–188.

Poulson, M., Neufeld, M. Y., Dechert, T., Allee, L., and Kenzik, K. M. (2021). Historic redlining, structural racism, and firearm violence: a structural equation modeling approach. *The Lancet Regional Health-Americas*, 3:100052.

Ratcliffe, J. (2010). Crime mapping: spatial and temporal challenges. *Handbook of quantitative criminology*, pages 5–24.

Ratcliffe, J. H. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science*, 18(1):61–72.

Rose, J. D. (2011). The incredible holc? mortgage relief during the great depression. *Journal of Money, Credit and Banking*, 43(6):1073–1107.

Ruggles, S., Fitch, C. A., Goeken, R., Hacker, J. D., Nelson, M. A., Roberts, E., Schouweiler, M., and Sobek, M. (2021). Ipums ancestry full count data: Version 3.0 [dataset]. *Minneapolis, MN: IPUMS*, 10:D010.

Ryan, H. (2018). Were the holc residential security maps the origin of redlining? *Oregon Historical Quarterly*, 119(4):538–543.

Sampson, R. J. and Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American journal of sociology*, 94(4):774–802.

Sharkey, P. (2008). The intergenerational transmission of context. *American Journal of Sociology*, 113(4):931–969.

Sharkey, P., Besbris, M., and Friedson, M. (2016). Poverty and crime.

Shertzer, A., Twinam, T., and Walsh, R. P. (2018). Zoning and the economic geography of cities. *Journal of Urban Economics*, 105:20–39.

Simonsohn, U., Simmons, J. P., and Nelson, L. D. (2020). Specification curve analysis. *Nature Human Behaviour*, 4(11):1208–1214.

Sood, A., Speagle, W., and Ehrman-Solberg, K. (2019). Long shadow of racial discrimination: Evidence from housing covenants of minneapolis. *Available at SSRN 3468520*.

Von Hoffman, A. (2008). The lost history of urban renewal. *Journal of Urbanism*, 1(3):281–301.

Weisburd, D., Bushway, S., Lum, C., and Yang, S.-M. (2004). Trajectories of crime at places: A longitudinal study of street segments in the city of seattle. *Criminology*, 42(2):283–322.

Wilson, W. J. (2003). Race, class and urban poverty: A rejoinder. *Ethnic & Racial Studies*, 26(6):1096–1114.

Wilson, W. J. (2008). The political and economic forces shaping concentrated poverty. *Political Science Quarterly*, 123(4):555–571.

Woods, L. L. (2012). The federal home loan bank board, redlining, and the national proliferation of racial lending discrimination, 1921–1950. *Journal of Urban History*, 38(6):1036–1059.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wright, C. D. (1893). The relation of economic conditions to the causes of crime: " the destruction of the poor is their poverty."-proverbs x. 15. *The ANNALS of the American Academy of Political and Social Science*, 3(6):96–116.

Xie, M. and Baumer, E. P. (2019). Crime victims' decisions to call the police: Past research and new directions. *Annual Review of Criminology*, 2(1):217–240.

Table 1: Descriptive statistics by HOLC grade, hexagonal data

	Mean (std. dev.)			
	Green (Grade A)	Blue (Grade B)	Yellow (Grade C)	Red (Grade D)
A. Present-day data (2014-2019)				
Crimes				
Index crimes	9.1 (22.3)	15.8 (28.5)	21.5 (38.0)	22.3 (40.4)
Violent	1.0 (3.3)	2.8 (7.7)	5.1 (10.7)	5.7 (11.1)
Murder	0.02 (0.1)	0.05 (0.3)	0.1 (0.4)	0.1 (0.4)
Robbery	0.4 (1.8)	1.2 (3.5)	2.0 (4.8)	2.1 (5.0)
Aggravated assault	0.5 (2.0)	1.6 (4.8)	3.0 (6.7)	3.4 (6.9)
Property	8.1 (20.3)	13.0 (23.6)	16.4 (31.3)	16.6 (33.4)
Burglary	1.8 (3.3)	3.0 (5.0)	3.8 (6.1)	3.6 (5.7)
Theft	5.4 (17.8)	8.1 (19.8)	9.9 (25.5)	10.4 (28.4)
Motor vehicle theft	1.0 (2.1)	1.9 (3.1)	2.7 (4.6)	2.6 (4.4)
Low level crimes	10.0 (31.1)	22.8 (57.3)	36.6 (87.7)	43.6 (107.4)
Dangerous weapons	0.1 (0.9)	0.5 (1.8)	0.9 (4.7)	1.1 (4.7)
Law Enforcement				
Index arrests	1.0 (23.2)	2.3 (27.2)	3.8 (35.2)	5.7 (55.2)
Violent	0.3 (4.7)	1.3 (18.5)	2.2 (22.9)	3.3 (33.8)
Murder	0.002 (0.1)	0.04 (1.3)	0.1 (1.7)	0.1 (2.2)
Robbery	0.1 (2.5)	0.5 (9.9)	0.7 (10.6)	1.0 (16.0)
Aggravated assault	0.2 (2.3)	0.8 (8.9)	1.4 (11.2)	2.1 (16.7)
Property	0.8 (19.1)	1.0 (10.8)	1.6 (14.0)	2.4 (25.4)
Burglary	0.2 (4.5)	0.3 (3.5)	0.4 (4.9)	0.7 (8.6)
Theft	0.6 (15.3)	0.6 (8.3)	0.9 (9.6)	1.5 (18.4)
Motor vehicle theft	0.03 (0.2)	0.1 (0.6)	0.2 (0.9)	0.3 (0.9)
Low level arrests	6.8 (47.3)	15.3 (80.8)	23.2 (115.6)	31.3 (147.7)
Dangerous weapons	0.1 (1.1)	0.6 (2.6)	1.0 (5.8)	1.4 (6.5)
Police stops	15.3 (68.1)	29.9 (143.0)	48.9 (215.9)	61.5 (298.6)
Calls for service	114.2 (403.9)	203.6 (599.6)	302.4 (731.1)	378.7 (829.3)
B. Historical data (1900-1930)				
Sociodemographics				
Black (%)	3.9 (14.9)	2.1 (11.6)	3.2 (15.1)	12.6 (29.8)
White (%)	95.8 (15.2)	97.8 (11.7)	96.7 (15.4)	87.1 (30.0)
Asian (%)	0.3 (3.3)	0.1 (2.0)	0.2 (2.9)	0.3 (4.0)
Native (%)	0.01 (0.3)	0.01 (0.5)	0.02 (0.8)	0.03 (1.3)
Foreign born (%)	12.0 (17.9)	15.4 (19.9)	17.9 (20.3)	19.6 (21.2)
Homeowner (%)	65.0 (40.5)	54.6 (40.3)	44.7 (38.2)	35.3 (35.9)
Illiterate (%)	0.7 (4.2)	0.8 (4.5)	1.4 (5.6)	3.0 (8.4)

Notes: Hexagonal mean and standard deviation within 500 meters from the closest border of a different HOLC grade. Pooled data from the 48 cities included in the study. A hexagonal area has size length of 100 meters and covers an area of 0.025 squared kilometers (around one city-block). Panel A includes present-day (2014-2019) public safety data. The data comes from incident level data from each city police department. Some cities do not report non-serious offenses and enforcement activity data. Index crimes include violent (murder, robbery, and aggravated assault) and property (burglary, theft, and motor vehicle theft) offenses. Dangerous weapons refer to criminal possession of a weapon and illegal discharge of a firearm. Low level crimes refer to all other offenses reported to law enforcement. Panel B includes historical (1900-1930) decennial census data. The data comes from pooled individual data from the 1900 1.2%, 1910 1.4%, 1920 1%, and 1930 5% decennial census samples. Homeownership refers that the inhabitants owned the housing unit, including those acquiring with a mortgage or other lending arrangement, even if repayment was not yet completed. Illiterate means that the person cannot read nor write.

Table 2: OLS regression discontinuity estimates on pre-HOLC maps census data

	Black	Foreign born	Homeowner	Illiterate
	(1)	(2)	(3)	(4)
A. Imbens and Kalyanaraman (2012) bandwidth				
Downgraded	0.02*	0.004	−0.01	0.004
	(0.01)	(0.01)	(0.02)	(0.003)
Change (%)	32.1%	2.2%	-2.9%	23.9%
Mean	0.05	0.18	0.45	0.02
Bandwidth	356.8	447.7	469.9	408.3
Observations	26,788	31,569	32,549	29,582
B. Calonico et al (2014) bandwidth				
Downgraded	0.01**	0.01	−0.03***	0.005***
	(0.005)	(0.01)	(0.01)	(0.001)
Change (%)	14.9%	4.2%	-7.2%	26.4%
Mean	0.06	0.18	0.44	0.02
Bandwidth	1,108.4	1,018.8	1,166.2	1,058.2
Observations	47,550	46,402	48,125	46,942

Notes: Regression discontinuity estimates of having a downgraded category relative to their neighboring border (a D-red grade relative a C-yellow, a C-yellow grade relative to B-blue, and a B-blue grade relative to an A-green). The outcomes measure pre-HOLC maps historical census data (1900-1930). The outcomes come from pooled individual data from the 1900 1.2%, 1910 1.4%, 1920 1%, and 1930 5% decennial census samples. Data aggregated to a hexagonal level using a tessellated grid. Each hexagon has a size length of 100 meters and covers an area of 0.025 squared kilometers (around one city-block). The model uses an ordinary least squares regression and a second-order polynomial function in the specification form. Robust standard errors in parentheses. The bandwidths selection follow Imbens and Kalyanaraman (2012) and Calonico et al. (2015) using a linear model. The bottom rows exhibit the percentage change (β/mean), followed the hexagonal mean between 1900-1930, the optimal bandwidth (in meters), and the number of observations. *p<0.1; **p<0.05; ***p<0.01.

Table 3: Poisson regression discontinuity estimates of HOLC maps on present-day crime

	Index	Violent	Property	Murder	Robbery	Assault	Burglary	Theft	Motor vehicle theft	Low level crimes	Weapons
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A. No controls											
Downgraded	0.12***	0.14***	0.11***	0.28***	0.16***	0.10***	0.10***	0.10***	0.11***	0.12***	0.17**
	(0.03)	(0.03)	(0.03)	(0.06)	(0.04)	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.08)
Change (%)	12.3%	14.9%	11.2%	32.0%	17.7%	10.1%	10.1%	11.0%	11.7%	12.6%	18.1%
Mean	18.4	4.2	14.2	0.1	1.6	2.5	3.2	8.7	2.2	31.0	0.8
Bandwidth	1,068.7	1,182.2	1,078.9	1,336.7	1,046.0	1,662.0	1,262.5	1,091.6	1,574.2	1,103.1	1,247.9
Observations	149,105	153,876	149,564	152,895	147,963	166,977	156,812	150,133	165,233	130,667	129,365
B. With historical census controls											
Downgraded	0.07***	0.09***	0.07***	0.21***	0.10***	0.08**	0.08***	0.06*	0.08***	0.07*	0.14**
	(0.02)	(0.03)	(0.03)	(0.06)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.07)
Change (%)	7.5%	9.3%	7.0%	23.3%	10.6%	7.9%	8.3%	6.0%	8.0%	7.1%	14.7%
Mean	23.0	5.4	17.6	0.1	2.1	3.2	3.9	10.9	2.7	40.5	1.0
Bandwidth	949.7	1,007.8	947.7	1,224.4	953.6	1,157.3	941.2	977.3	1,141.9	875.1	1,089.1
Observations	104,211	105,931	104,144	107,074	104,338	109,505	103,912	105,048	109,319	89,310	91,882

Notes: Regression discontinuity estimates on the crime effects of having a downgraded category relative to its neighboring border (a D-red grade relative to a C-yellow, a C-yellow grade relative to B-blue, and a B-blue grade relative to an A-green). The data comes from incident level records between 2014-2019 from each city police department, aggregated to a hexagonal level using a tessellated grid. Each hexagon has a size length of 100 meters and covers an area of 0.025 squared kilometers (around one city-block). The model uses a Poisson regression and a second-order polynomial function in the specification form. Robust standard errors in parentheses. Panel A presents the estimates using the spatial regression discontinuity as Equation (1). Panel B adds preintervention (1900-1930) sociodemographic controls (proportion of Black, White, homeowners, and illiterate population). Due to missing historical data, the sample size decreases in Panel B. The bandwidth selection follows Calonico et al. (2015) using a linear model. Index crimes include violent (murder, robbery, and aggravated assault) and property (burglary, theft, and motor vehicle theft) offenses. Dangerous weapons refer to criminal possession of a weapon and illegal firearm discharges. Low level crimes refer to all other offenses reported to law enforcement. The bottom rows exhibit the percentage change (incidence rate ratio - 1 = exp(β)-1), followed by the hexagonal mean between 2014-2019, the optimal bandwidth (in meters), and the number of observations. *p<0.1; **p<0.05; ***p<0.01.

Table 4: Poisson regression discontinuity estimates of HOLC maps on police arrests

	Index	Violent	Property	Murder	Robbery	Assault	Burglary	Theft	Motor vehicle theft	Low level offenses	Weapons
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A. No controls											
Downgraded	0.14	0.14	0.16	0.48	0.24	0.07	0.22	0.15	0.19**	0.11	0.24**
	(0.21)	(0.22)	(0.21)	(0.37)	(0.32)	(0.18)	(0.23)	(0.24)	(0.08)	(0.10)	(0.12)
Change (%)	14.8%	14.7%	17.7%	62.1%	26.6%	7.4%	24.7%	16.5%	21.0%	12.0%	27.7%
Mean	3.5	2.0	1.5	0.1	0.6	1.3	0.4	0.9	0.2	20.3	0.9
Bandwidth	1,325.5	1,366.1	1,300.7	1,978.5	1,400.5	1,326.9	1,496.4	1,260.4	1,012.2	1,096.7	1,215.4
Observations	78,944	79,478	78,622	76,890	78,676	77,766	79,679	76,875	72,280	73,509	74,347
B. With historical census controls											
Downgraded	−0.04	−0.11	0.05	0.02	−0.19	−0.07	0.05	0.10	0.11	0.07	0.20**
	(0.19)	(0.22)	(0.20)	(0.39)	(0.30)	(0.17)	(0.24)	(0.24)	(0.08)	(0.09)	(0.10)
Change (%)	-4.0%	-10.7%	5.1%	1.9%	-17.1%	-6.4%	4.8%	10.4%	11.1%	7.1%	22.1%
Mean	4.6	2.7	1.9	0.1	0.9	1.8	0.5	1.2	0.3	27.0	1.2
Bandwidth	1,043.9	1,078.0	1,035.5	1,536.8	1,106.7	1,081.8	1,268.1	1,027.3	824.6	894.1	1,007.3
Observations	54,718	55,184	54,610	54,385	55,077	54,775	56,640	54,044	50,609	50,913	52,237

Notes: Regression discontinuity police arrests estimates of being inside a lower grade relative to its neighbor (a D-red grade relative to C-yellow, a C-yellow grade relative to B-blue, and a B-blue grade relative to A-green). The data comes from incident level records between 2014-2019 from each city police department, aggregated to a hexagonal level using a tessellated grid. Each hexagon has a size length of 100 meters and covers an area of 0.025 squared kilometers (around one city-block). The model uses a Poisson regression and a second-order polynomial function in the specification form. Robust standard errors in parentheses. Panel A presents the estimates using the spatial regression discontinuity as Equation (1). Panel B adds preintervention (1900-1930) sociodemographic controls (proportion of Black, White, homeowners, and illiterate population). Due to missing historial data, the sample size decreases in Panel B. The bandwidths selection follows Calonico et al. (2015) using a linear model. Index crimes include violent (murder, robbery, and aggravated assault) and property (burglary, theft, and motor vehicle theft) offenses. Dangerous weapons refer to criminal possession of a weapon and illegal firearm discharges. Low level crimes refer to all other offenses reported to law enforcement. The bottom rows exhibit the percentage change (incidence rate ratio - 1 = exp(β)-1), followed by the hexagonal mean between 2014-2019, the optimal bandwidth (in meters), and the number of observations. *p<0.1; **p<0.05; ***p<0.01.

Table 5: Poisson regression discontinuity estimates of HOLC maps on police stops and calls for service

	Police stops	Calls for service
	(1)	(2)
A. No controls		
Downgraded	0.16	−0.04
	(0.10)	(0.06)
Change (%)	17.5%	-4.2%
Mean	43.4	261.8
Bandwidth	1,023.1	934.9
Observations	70,159	53,726
B. With historical census controls		
Downgraded	0.15^{*}	-0.10^{*}
	(0.09)	(0.06)
Change (%)	16.2%	-9.5%
Mean	52.7	331.0
Bandwidth	879.1	732.0
Observations	53,454	37,268

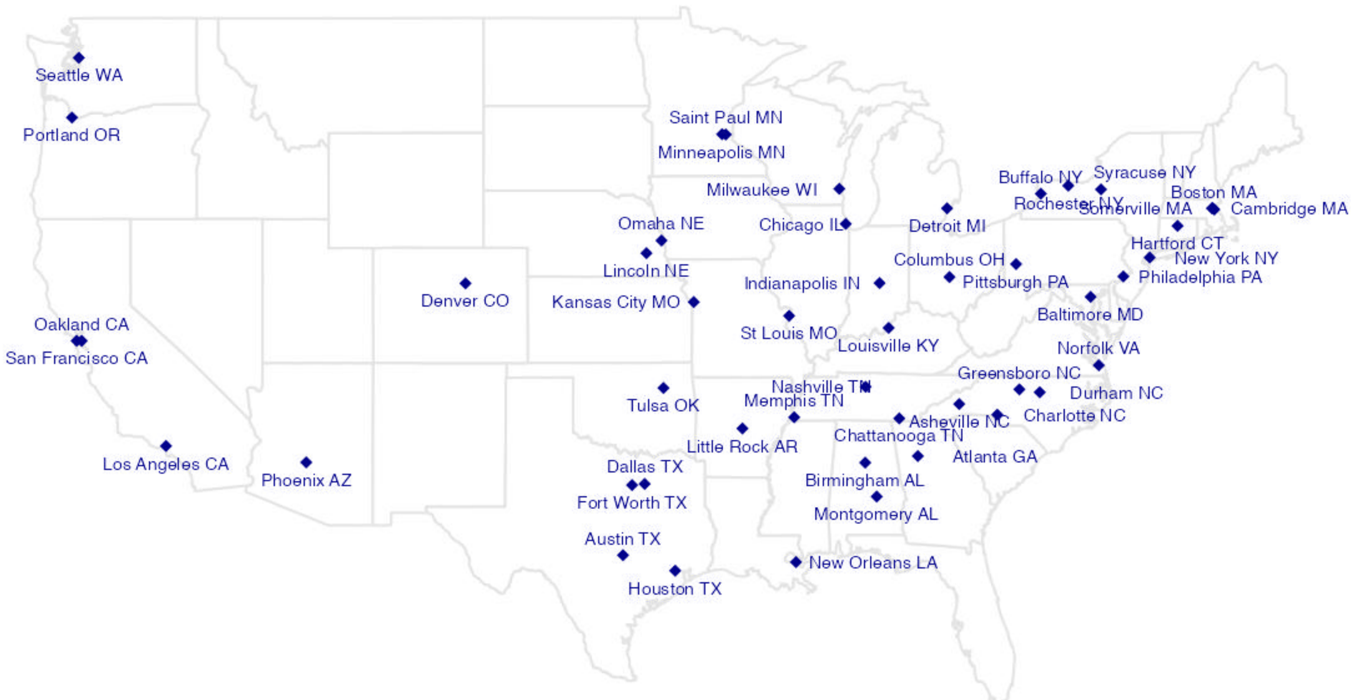
Notes: Regression discontinuity police stops and calls for service estimates of being inside a lower grade relative to its neighbor (a D-red grade relative to C-yellow, a C-yellow grade relative to B-blue, and a B-blue grade relative to A-green). The data comes from incident level records between 2014-2019 from each city police department, aggregated to a hexagonal level using a tessellated grid. Each hexagon has a size length of 100 meters and covers an area of 0.025 squared kilometers (around one city-block). The model uses a Poisson regression and a second-order polynomial function in the specification form. Robust standard errors in parentheses. Panel A presents the estimates using the spatial regression discontinuity as Equation (1). Panel B adds preintervention (1900-1930) sociodemographic controls (proportion of Black, White, homeowners, and illiterate population). Due to missing historial data, the sample size decreases in Panel B. The bandwidths selection follows Calonico et al. (2015) using a linear model. The bottom rows exhibit the percentage change (incidence rate ratio - 1 = exp(β)-1), followed by the hexagonal mean between 2014-2019, the optimal bandwidth (in meters), and the number of observations. *p<0.1; **p<0.05; ***p<0.01.

Table 6: Poisson difference-in-discontinuities estimates of HOLC maps on present-day crime by baseline levels

	Index	Violent	Property	Murder	Robbery	Assault	Burglary	Theft	Motor vehicle theft	Low level crimes	Weapons
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A. Homeowner											
Downgraded*Homeowner	0.08	0.16***	0.06	0.21*	0.16**	0.12**	0.06	0.05	0.09**	0.06	0.16
	(0.05)	(0.06)	(0.05)	(0.11)	(0.07)	(0.06)	(0.04)	(0.07)	(0.04)	(0.07)	(0.11)
B. Foreign born											
Downgraded*Foreign-born	0.14	0.13	0.11	0.04	0.09	0.19	0.20***	0.03	0.25***	0.02	0.13
	(0.09)	(0.12)	(0.10)	(0.21)	(0.13)	(0.12)	(0.08)	(0.13)	(0.08)	(0.13)	(0.18)
C. Non white											
Downgraded*Non-white	−0.04	0.08	−0.07	0.15	0.03	0.13	−0.02	−0.08	0.02	−0.03	−0.08
	(0.07)	(0.08)	(0.07)	(0.15)	(0.09)	(0.08)	(0.06)	(0.09)	(0.07)	(0.10)	(0.15)
Mean	23.0	5.4	17.6	0.1	2.1	3.2	3.9	10.9	2.7	40.5	1.0
Bandwidth	949.7	1,007.8	947.7	1,224.4	953.6	1,157.3	941.2	977.3	1,141.9	875.1	1,089.1
Observations	104,211	105,931	104,144	107,074	104,338	109,505	103,912	105,048	109,319	89,310	91,882

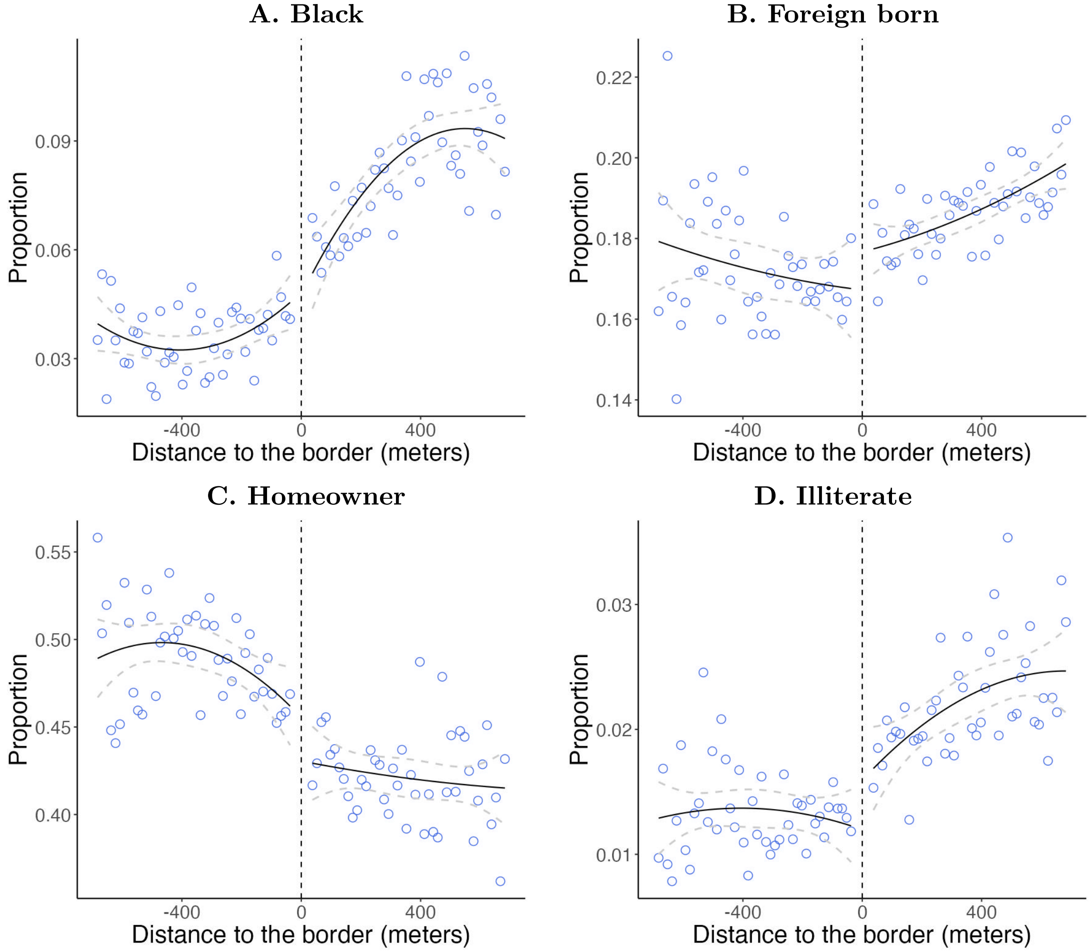
Notes: Difference-in-discontinuities estimates on the crime effects of having a downgraded category relative to its neighboring border (a D-red grade relative to a C-yellow, a C-yellow grade relative to B-blue, and a B-blue grade relative to an A-green), interacting the parameters with the relevant pre-treatment heterogeneity dimension. Specifically, the model uses the following specification: $y_i = \alpha_0 + \alpha_1 f(r_i) + D_i(\alpha_2 + \alpha_3 f(r_i)) + H_i[\alpha_4 + \alpha_5 f(r_i) + D_i(\beta_1 + \alpha_6 f(r_i))] + e_i$, where H_i is the pre-HOLC map heterogeneity variable measured as an indicator variable. The table presents the coefficient β_1. Panel A, homeowner, refers to having more than 50 percent of the residents being homeowners. Panel B, foreign-born, refers to having more than 50 percent of the residents being born abroad. Panel C, non-white, refers to having a positive percent of the population be non-white. The heteroneity variabels comes the 1900-1930 census sociodemographic data. The crime data comes from incident level records between 2014-2019 from each city police department, aggregated to a hexagonal level using a tessellated grid. Each hexagon has a size length of 100 meters and covers an area of 0.025 squared kilometers (around one city-block). The model uses a Poisson regression and a second-order polynomial function in the specification form. Robust standard errors in parentheses. The bandwidth selection follows Calonico et al. (2015) using a linear model. Index crimes include violent (murder, robbery, and aggravated assault) and property (burglary, theft, and motor vehicle theft) offenses. Dangerous weapons refer to criminal possession of a weapon and illegal firearm discharges. Low level crimes refer to all other offenses reported to law enforcement. The bottom rows exhibit the hexagonal mean between 2014-2019, the optimal bandwidth (in meters), and the number of observations. *p<0.1; **p<0.05; ***p<0.01.

Figure 1: Cities included in the analysis



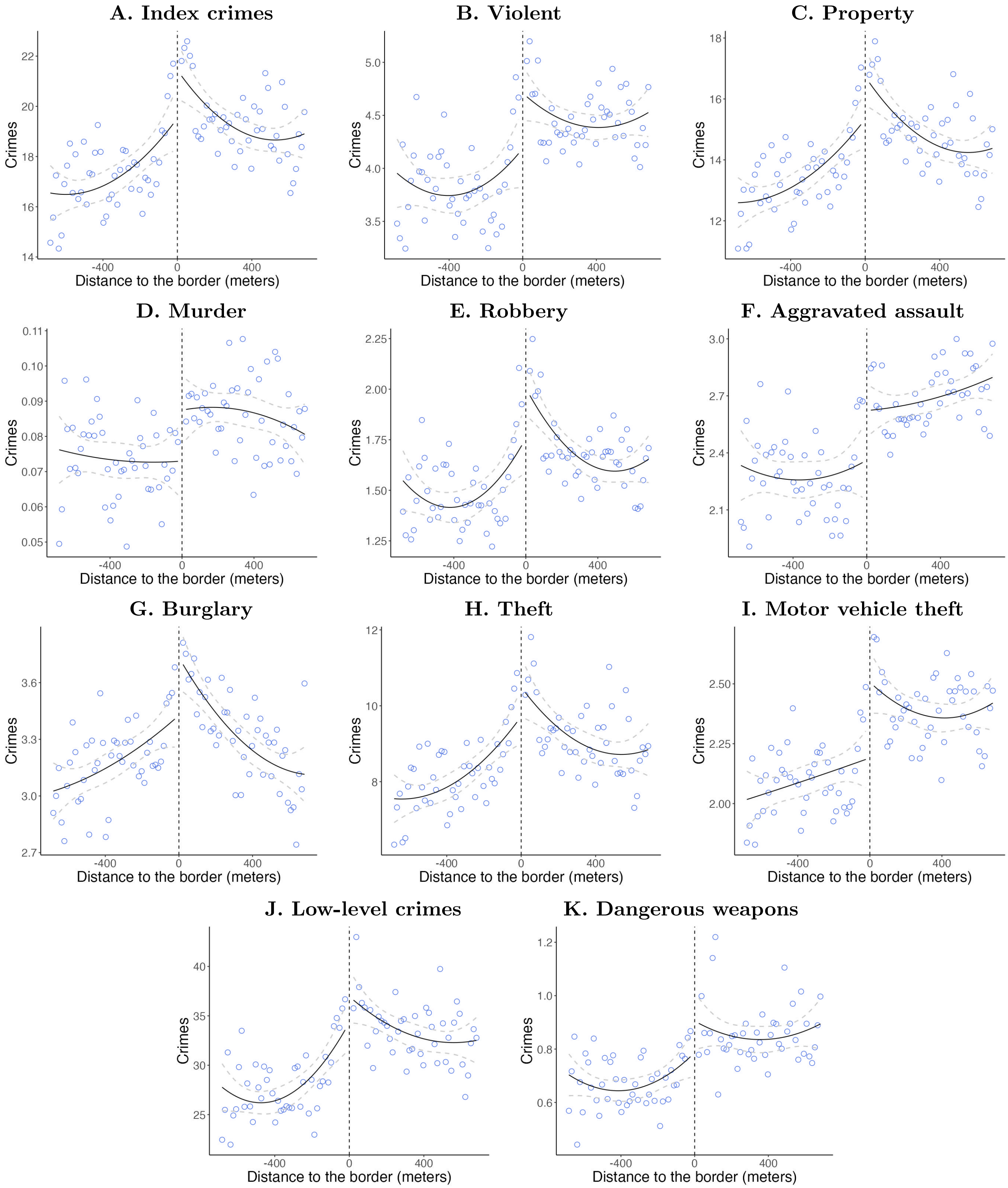
Notes: The map shows the location of the 48 US cities included in this research. The cities have a HOLC map and georeferenced public crime data from 2014 to 2019.

Figure 2: Regression discontinuity of HOLC maps on 1900-1930s sociodemographics



Notes: Mean proportions per hexagon. The data comes from pooled individual records from the 1900 1.2%, 1910 1.4%, 1920 1%, and 1930 5% decennial census samples, aggregated to a tessellated hexagonal grid (the unit of analysis). Hexagonal means using bins of 15 meters, second-order polynomials (solid line), and 95 percent confidence intervals (dash lines) around a 700-meter distance of the closest border of a different HOLC grade. A positive distance means being in a lower grade relative to its neighbor (a red grade relative to yellow, a yellow grade relative to blue, and a blue grade relative to green).

Figure 3: Regression discontinuity estimates of HOLC maps on present-day reported crime

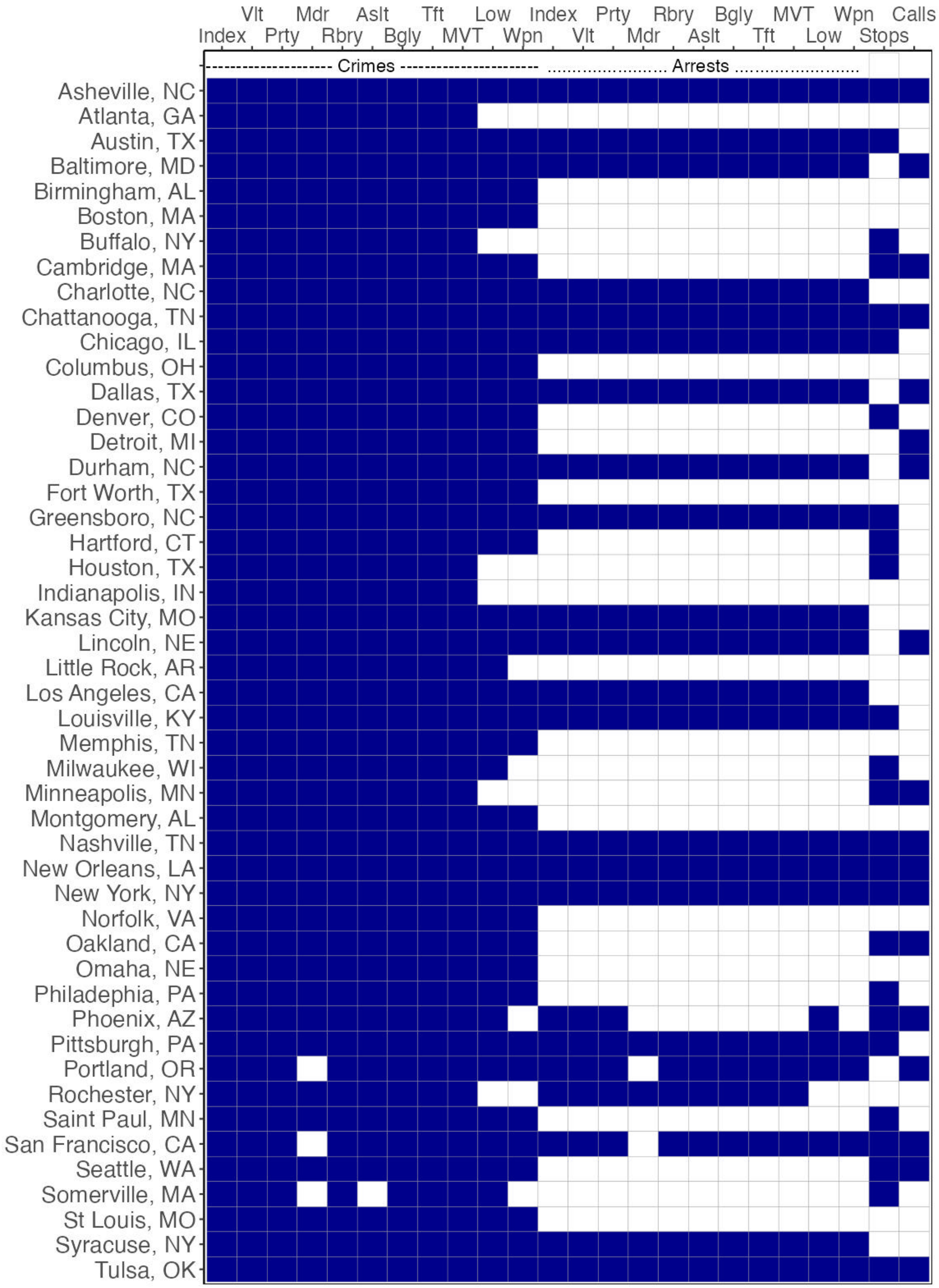


Notes: Mean crimes per hexagon between 2014 and 2019, using bins of 15 meters, second-order polynomials (solid line), and 95 percent confidence intervals (dash lines) around a 700-meter distance of the closest border of a different HOLC grade. A positive distance means being in a lower grade relative to its neighbor (a red grade relative to yellow, a yellow grade relative to blue, and a blue grade relative to green). Index crimes include violent (murder, robbery, and aggravated assault) and property (burglary, theft, and motor vehicle theft) offenses. Dangerous weapons refer to criminal possession of a weapon and illegal firearm discharges. Non-serious crimes refer to all other offenses reported to law enforcement.

ONLINE APPENDIX

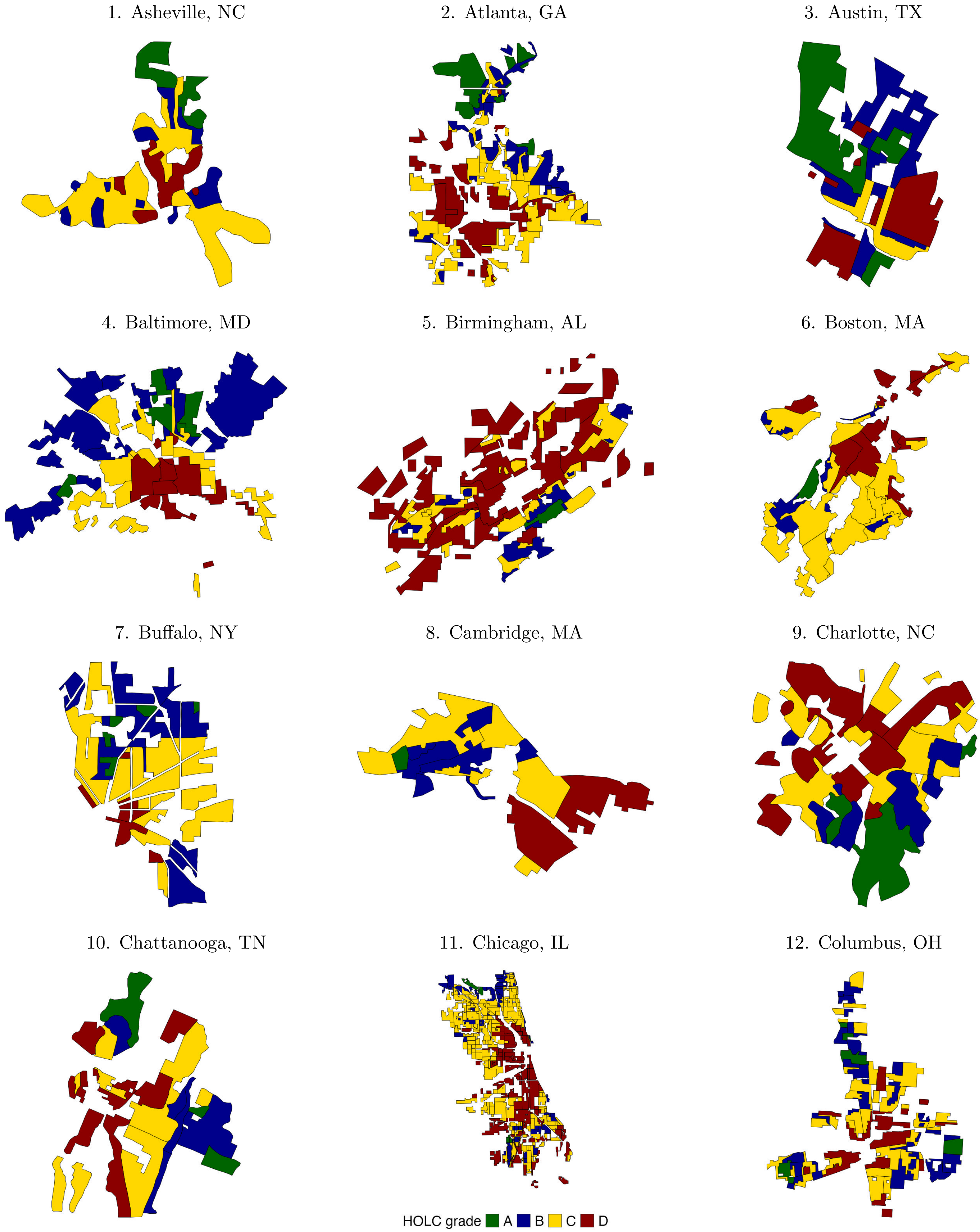
A Appendix: City-specific data processing

Figure A.1: Data availability by city

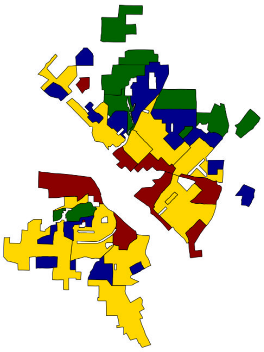


Notes: The figure shows the available data by city. The abbreviations stand for the following: violent (vlt), property (prty), murder (mdr), robbery (rbry), burglary (bgly), theft (TFT), motor vehicle theft (mvt), low-level offenses (low), weapon (wpn).

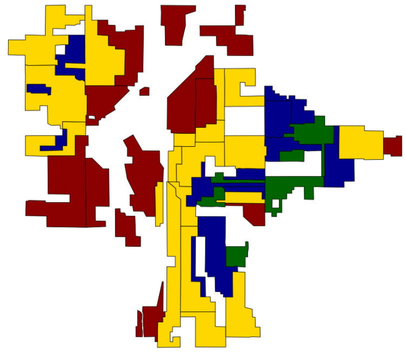
Figure A.2: HOLC maps by city



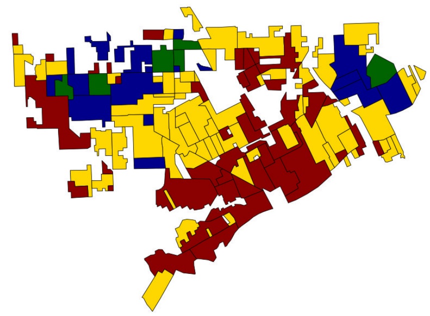
13. Dallas, TX



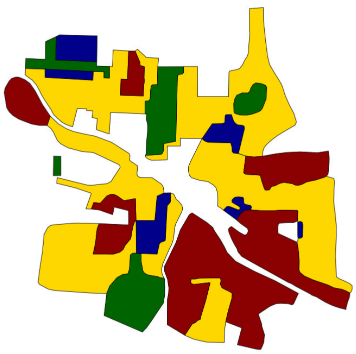
14. Denver, CO



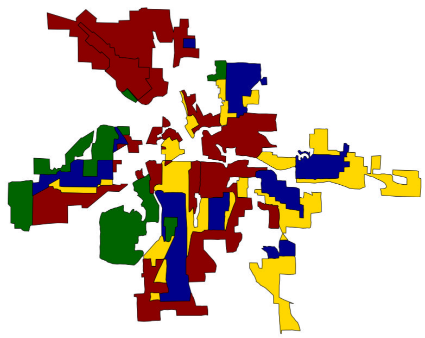
15. Detroit, MI



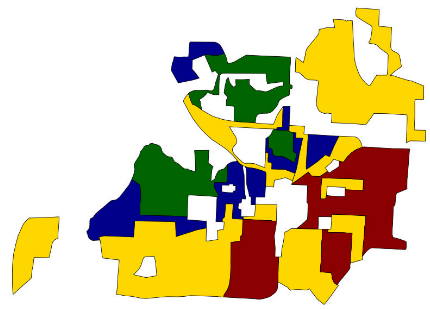
16. Durham, NC



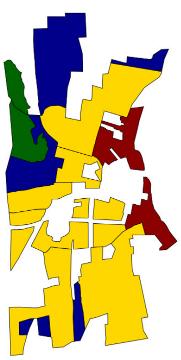
17. Fort Worth, TX



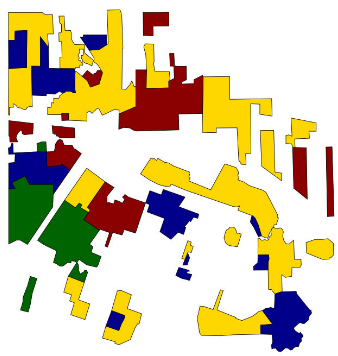
18. Greensboro, NC



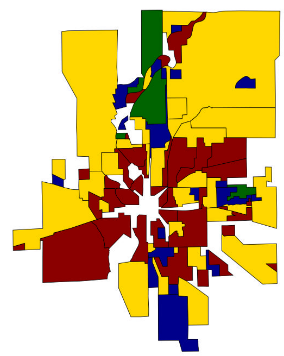
19. Hartford, CT



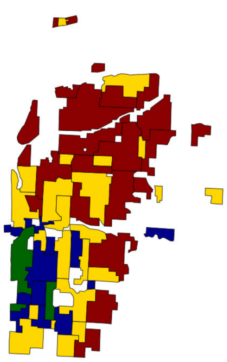
20. Houston, TX



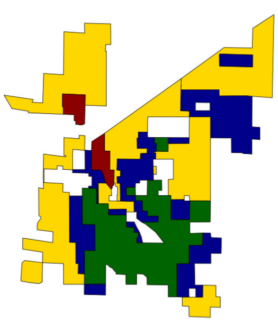
21. Indianapolis, IN



22. Kansas City, MO



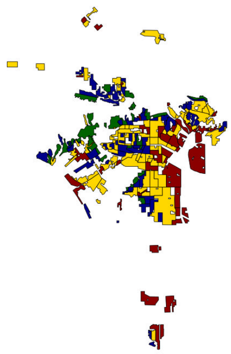
23. Lincoln, NE



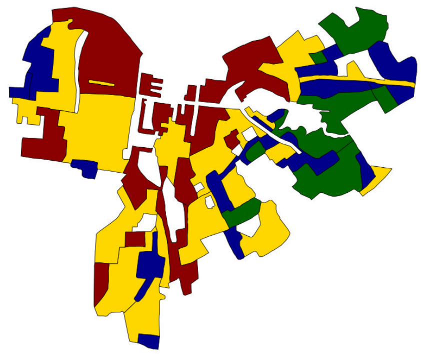
24. Little Rock, AR

HOLC grade A B C D

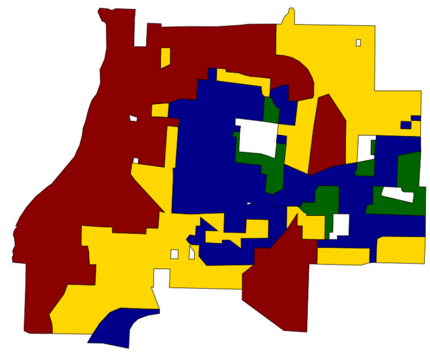
25. Los Angeles, CA



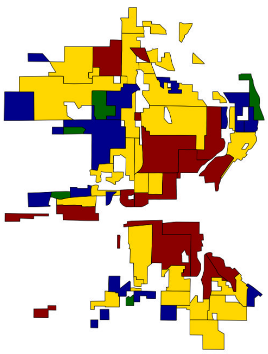
26. Louisville, KY



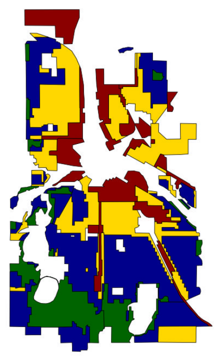
27. Memphis, TN



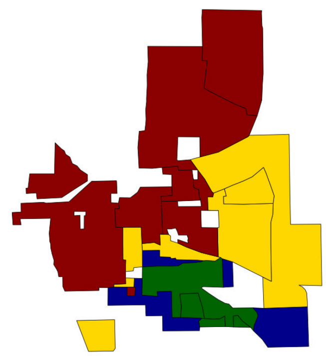
28. Milwaukee, WI



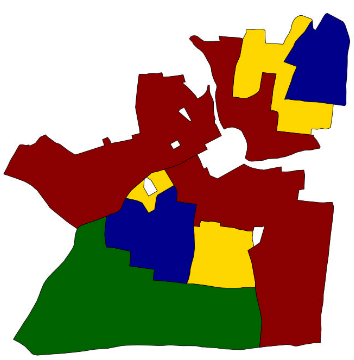
29. Minneapolis, MN



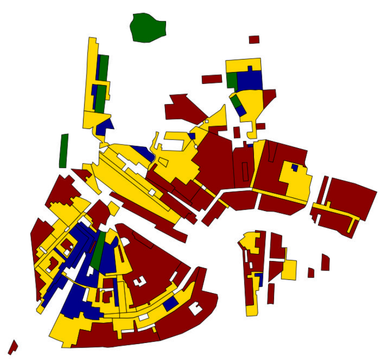
30. Montgomery, AL



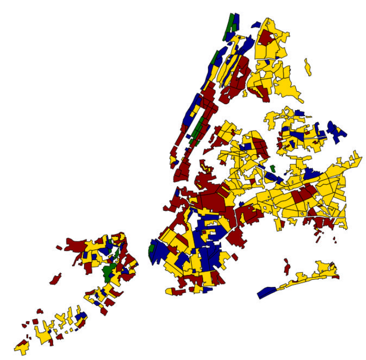
31. Nashville, TN



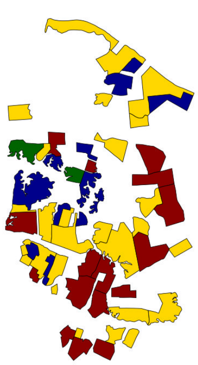
32. New Orleans, LA



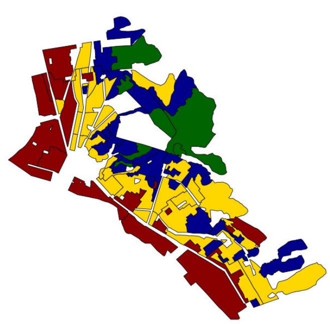
33. New York, NY



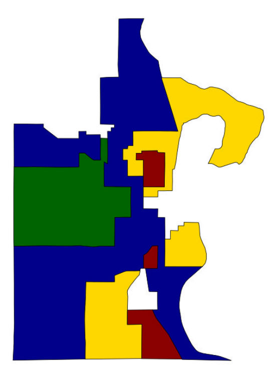
34. Norfolk, VA



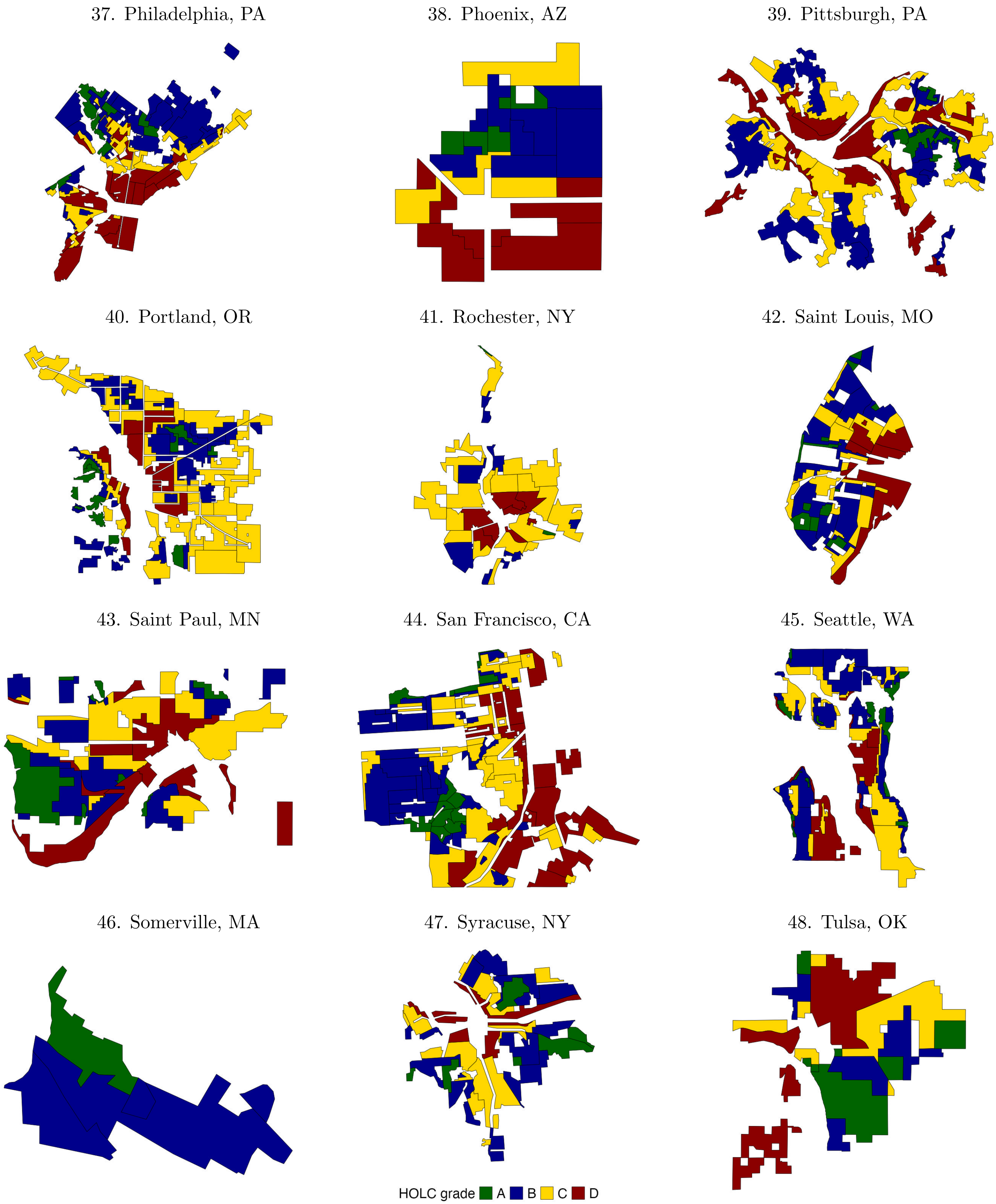
35. Oakland, CA



36. Omaha, NE



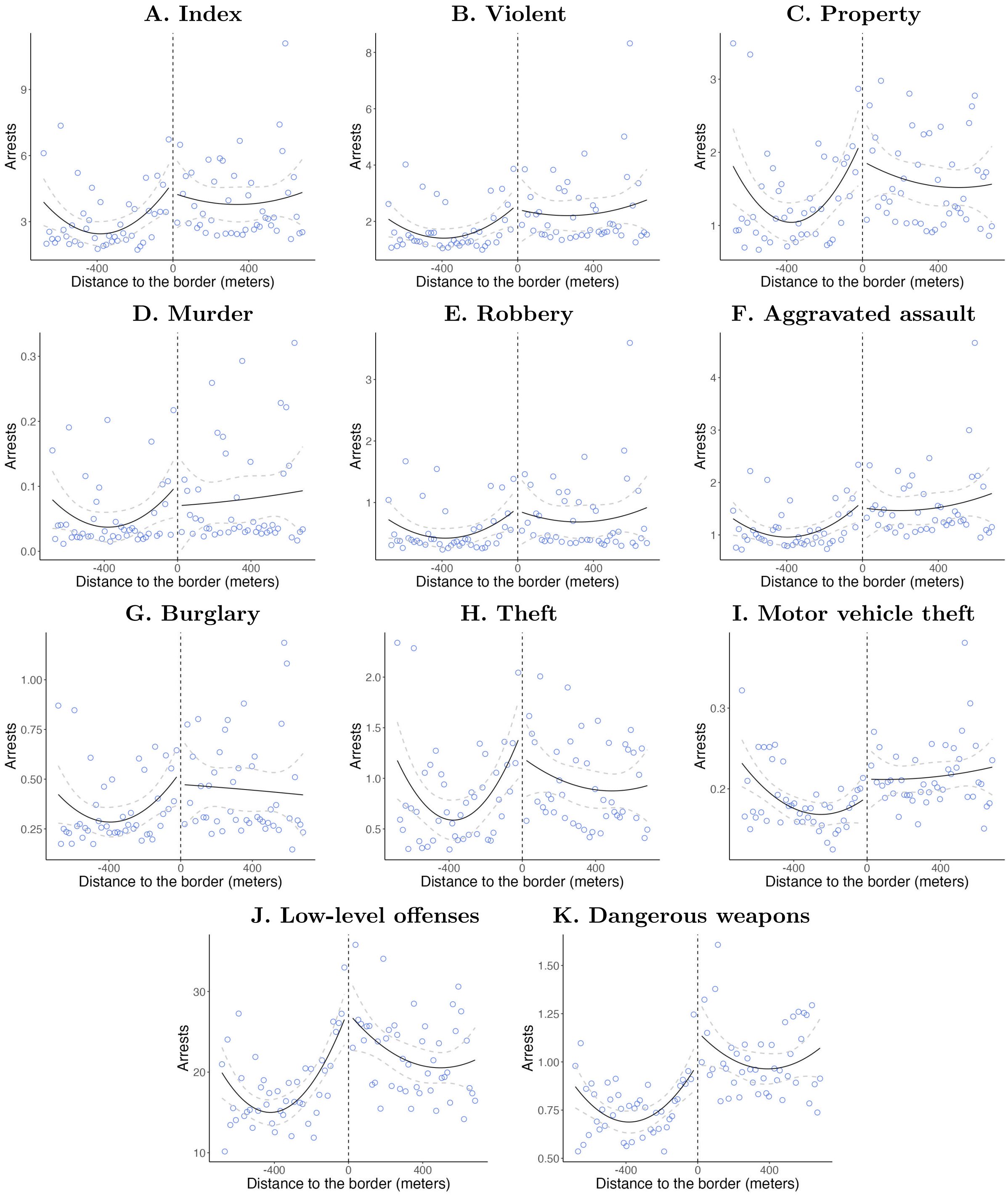
HOLC grade A B C D



Notes: The Los Angeles and Norfolk HOLC maps only include the area inside their respective cities, as their police departments and, hence, the public safety availability only cover such jurisdiction.

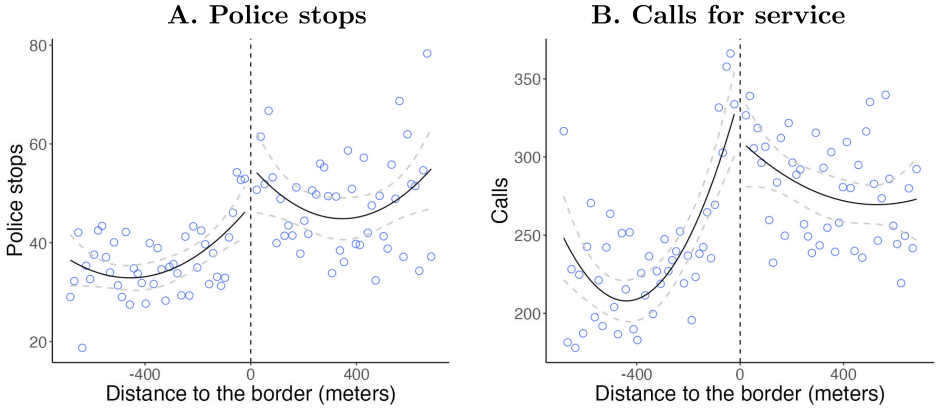
B Appendix: Additional law enforcement outcomes

Figure B.1: Regression discontinuity of HOLC maps on police arrests



Notes: Mean number of arrests per hexagon between 2014 and 2019, using bins of 50 meters, second-order polynomials (solid line), and 95 percent confidence intervals (dash lines) around a 500 meters distance of the closest border of a different HOLC grade. A positive distance means being inside a lower grade relative to its neighbor (a red grade relative to yellow, a yellow grade relative to blue, and a blue grade relative to green). Index crimes include violent (murder, robbery, and aggravated assault) and property (burglary, theft, and motor vehicle theft) offenses. Dangerous weapons refer to criminal possession of a weapon and illegal discharge of a firearm. Low-level refers to all other offenses reported to law enforcement.

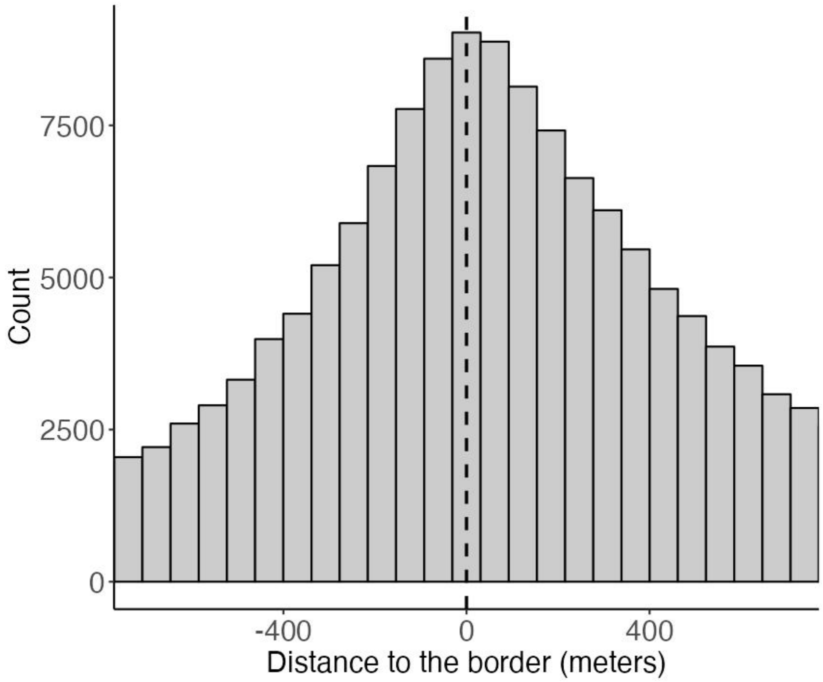
Figure B.2: Regression discontinuity of HOLC maps on police stops and calls for service



Notes: Mean number of arrests per hexagon between 2014 and 2019, using bins of 50 meters, second-order polynomials (solid line), and 95 percent confidence intervals (dash lines) around a 500 meters distance of the closest border of a different HOLC grade. A positive distance means being inside a lower grade relative to its neighbor (a red grade relative to yellow, a yellow grade relative to blue, and a blue grade relative to green).

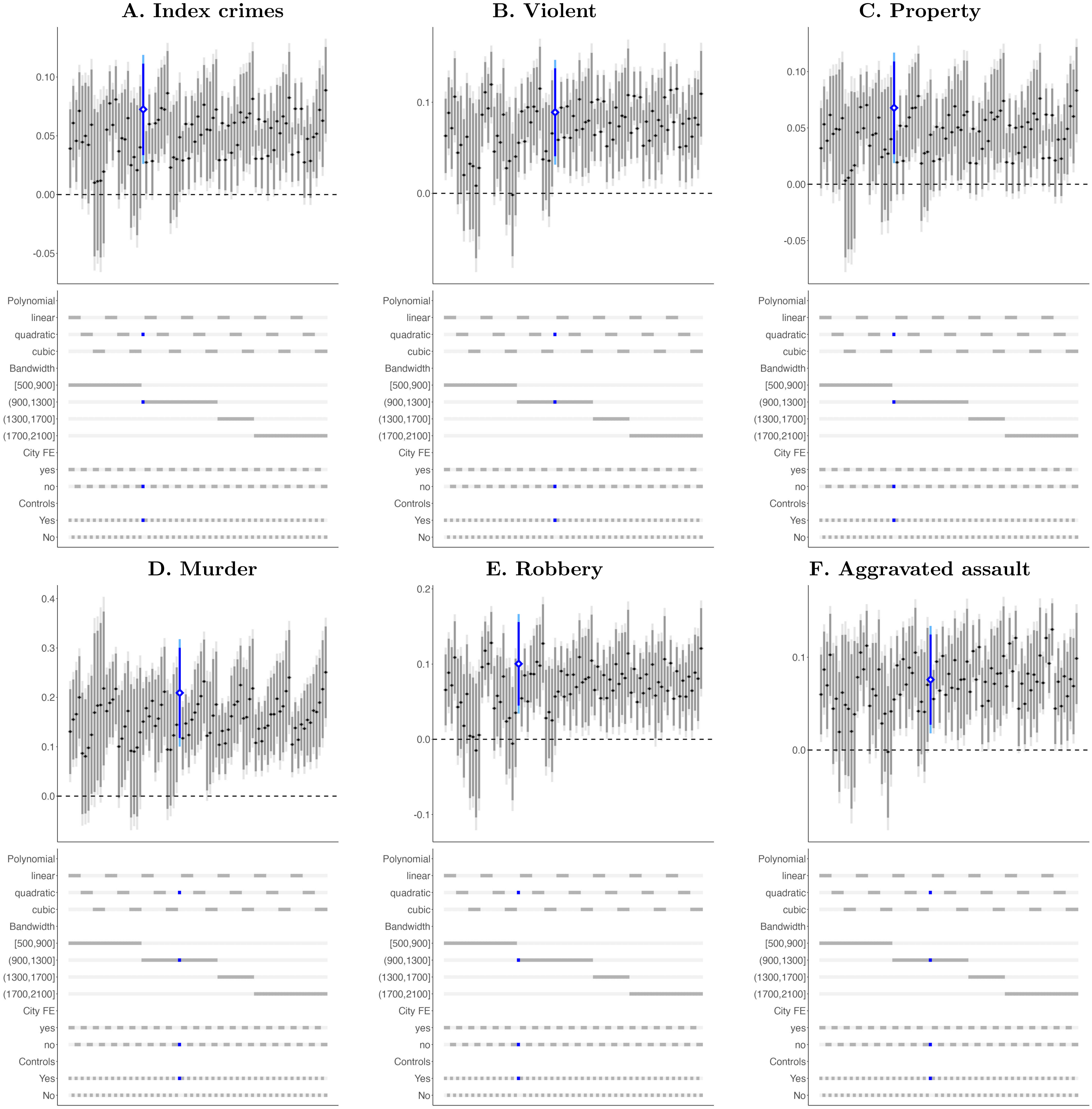
C Appendix: Identifying assumptions and robustness checks

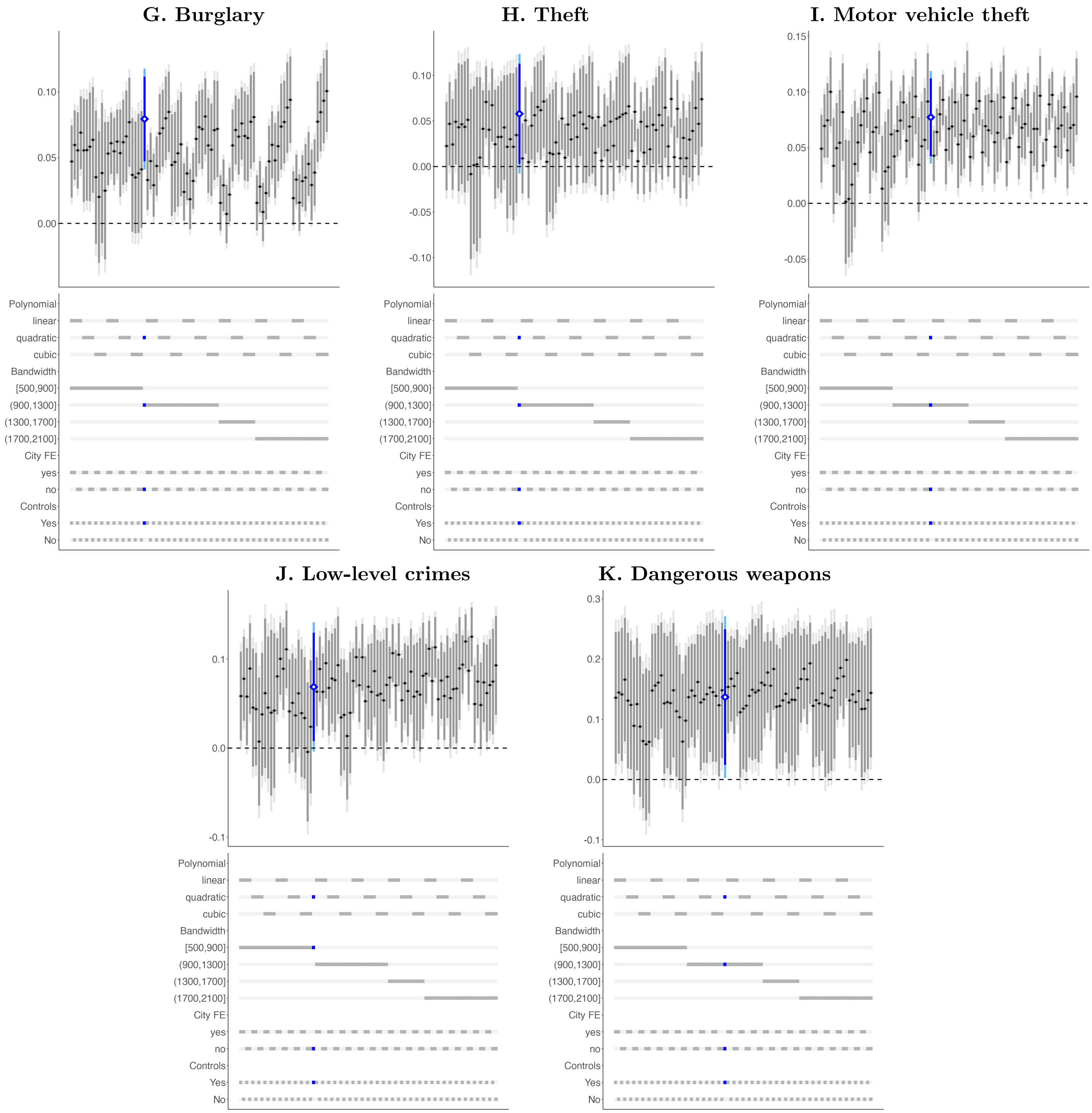
Figure C.1: Running variable distribution



Notes: Distribution of hexagons by distance to the closest border within 700 meters of the border. A positive distance means being inside a lower grade relative to its neighbor (a red grade relative to yellow, a yellow grade relative to blue, and a blue grade relative to green). The pvalues of the manipulation tests are 0.56 and 0.94 using the McCrary (2008) and Cattaneo et al. (2018) approaches.

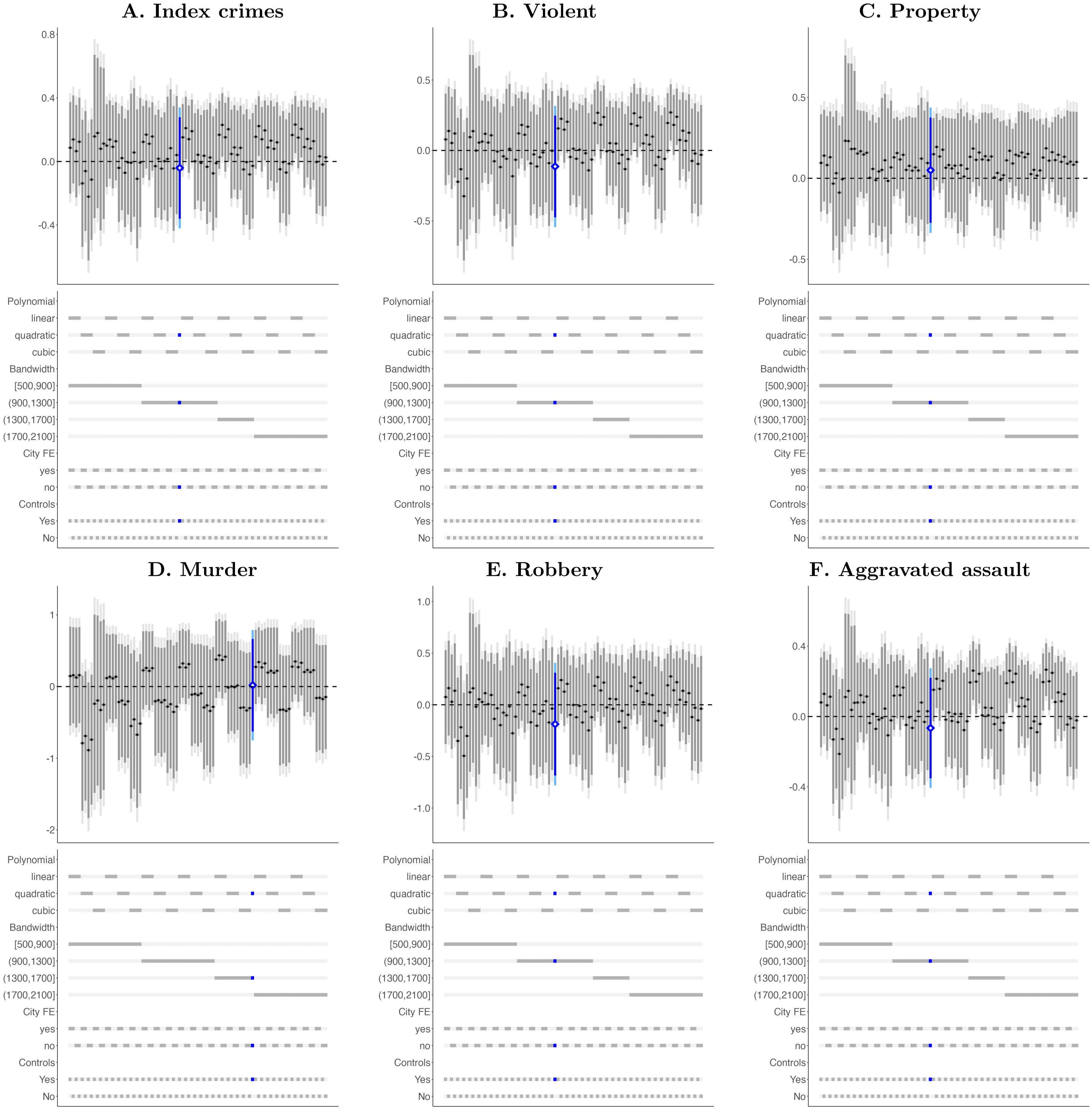
Figure C.2: Poisson regression discontinuity of HOLCS maps on present-day crime, alternative specifications



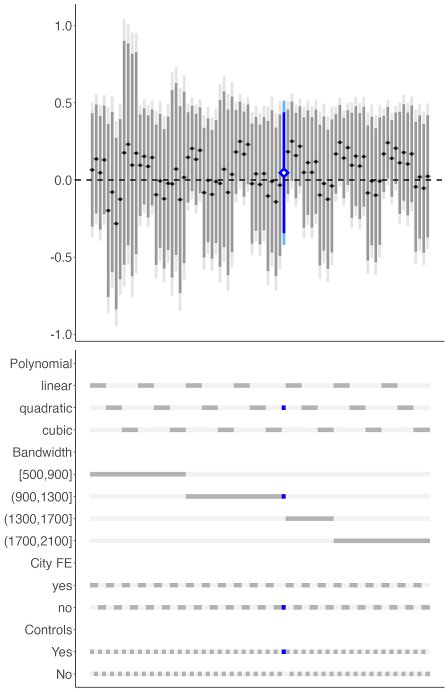


Notes: Regression discontinuity estimates using a Poisson regression for different bandwidths (from 500 to 2,000 meters), polynomial functions (1st, 2nd, and 3rd), fixed effects (city fixed effects or without them), and controlling for pre-HOLC maps sociodemographics (with and without controls). The main specification, highlighted in blue, is estimated using Calonico et al. (2015) bias-corrected bandwidth.

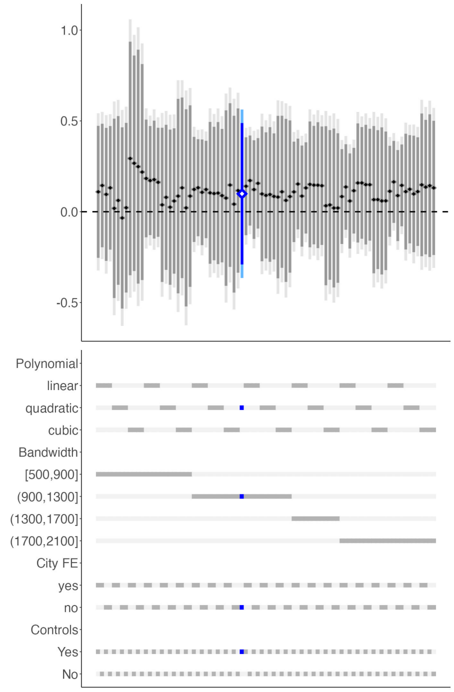
Figure C.3: Poisson regression discontinuity of HOLC maps on present-day police arrests, alternative specifications



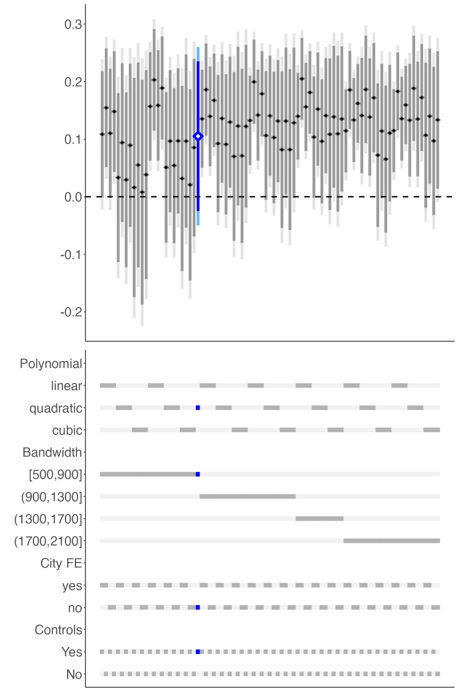
G. Burglary



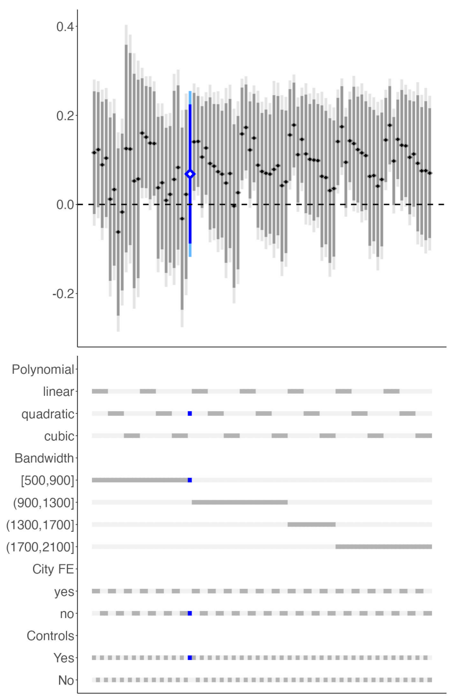
H. Theft



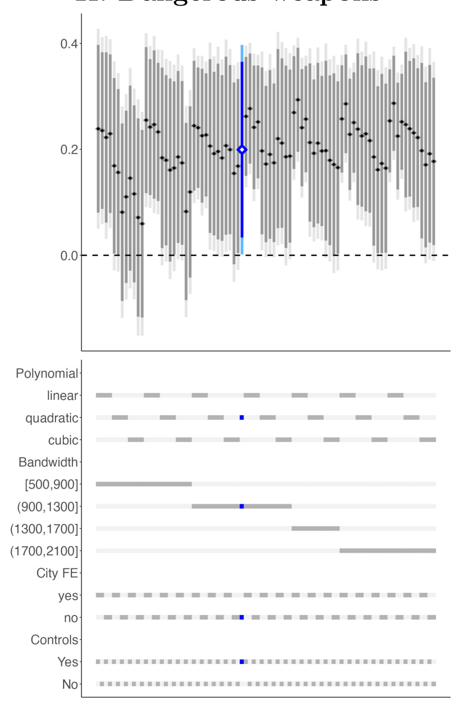
I. Motor vehicle theft



J. Low-level crimes

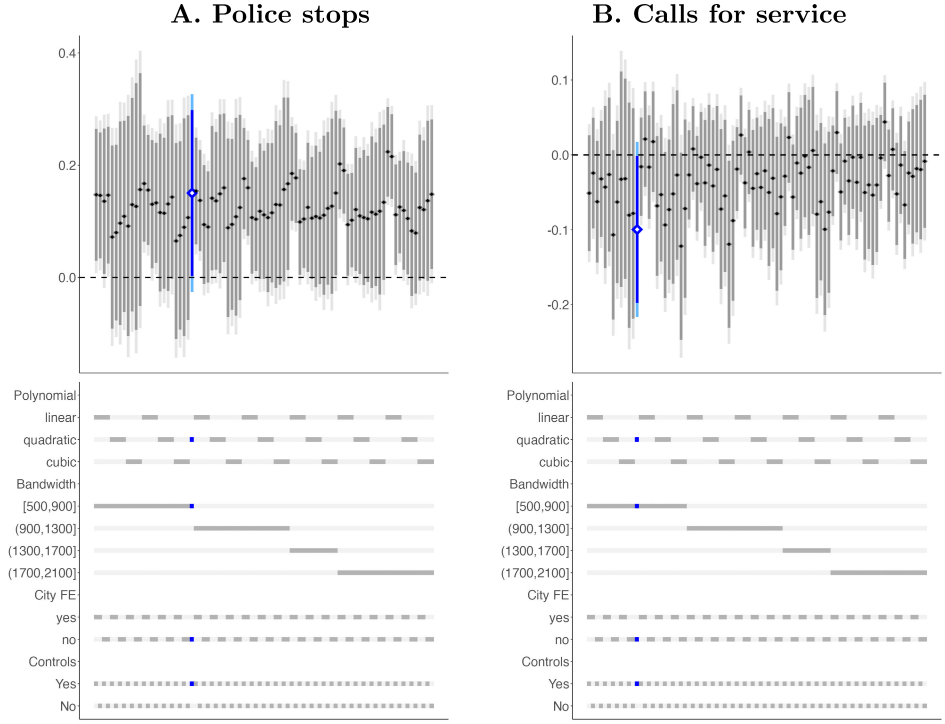


K. Dangerous weapons



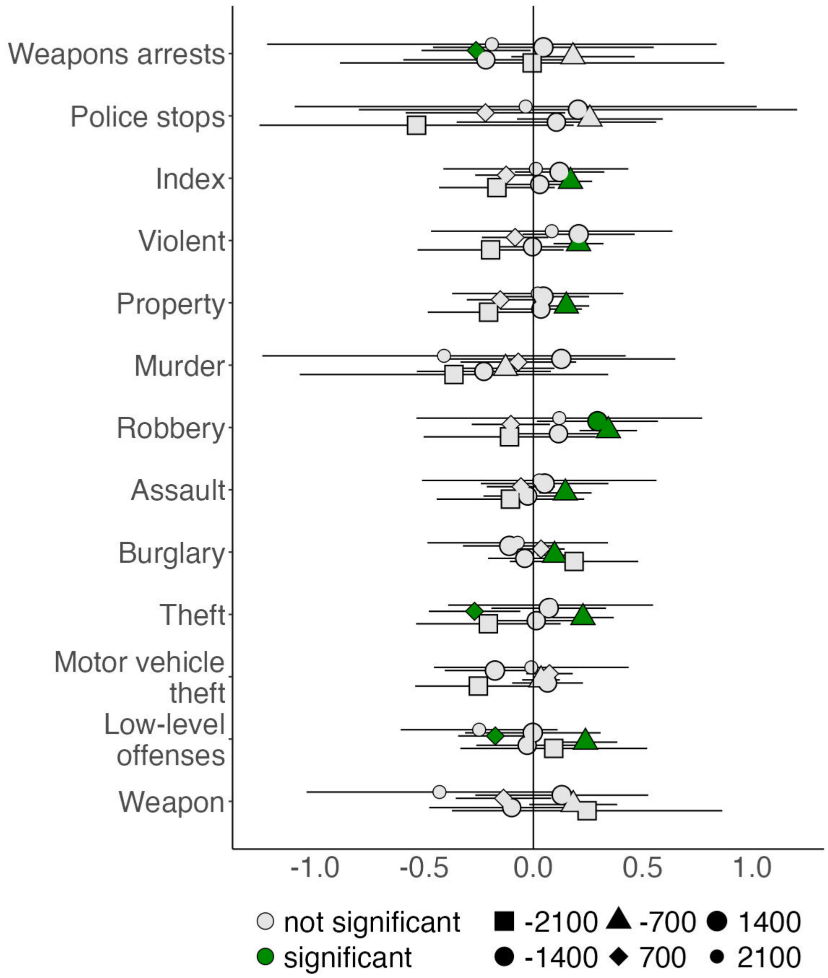
Notes: Regression discontinuity estimates using a Poisson regression for different bandwidths (from 500 to 2,000 meters), polynomial functions (1st, 2nd, and 3rd), fixed effects (city fixed effects or without them), and controlling for pre-HOLC maps sociodemographics (with and without controls). The main specification, highlighted in blue, is estimated using Calonico et al. (2015) bias-corrected bandwidth.

Figure C.4: Poisson regression discontinuity of HOLC maps on present-day police stops and calls for service, alternative specifications



Notes: Regression discontinuity estimates using a Poisson regression for different bandwidths (from 500 to 2,000 meters), polynomial functions (1st, 2nd, and 3rd), fixed effects (city fixed effects or without them), and controlling for pre-HOLC maps sociodemographics (with and without controls). The main specification, highlighted in blue, is estimated using Calonico et al. (2015) bias-corrected bandwidth.

Figure C.5: Placebo test on the Poisson regression discontinuity estimates at different cutoff values



Notes: Poisson regression discontinuity crime, police stop, and arrest estimates on the effects of being in a lower grade relative to its neighbor (a red grade relative to yellow, a yellow grade relative to blue, and a blue grade relative to green). The specification uses a second-order polynomial function. The placebo test shifts the threshold of the closest distance to the border of a different HOLC grade area (running variable) ranging from -2,100 to +2,100 meters in 700 meters increments and estimates the optimal bandwidth following Calonico et al. (2015) using a second-order polynomial function. The 95 percent confidence intervals and their statistical significance are marked in the figure.