Race, Risk, and the Emergence of Federal Redlining[1]

Price V. Fishback (University of Arizona and NBER)
Jessica LaVoice (Bowdoin College)
Allison Shertzer (University of Pittsburgh and NBER)
Randall P. Walsh (University of Pittsburgh and NBER)

November 2020

Abstract: During the late 1930s, the Home Owners Loan Corporation (HOLC) created a series of maps designed to summarize spatial variation in the riskiness of mortgage lending in different neighborhoods. The HOLC maps, in conjunction with contemporaneous maps produced by the Federal Housing Agency (FHA), are at the center of debates regarding the long-run impacts of government-imposed redlining, particularly because black households were concentrated in the highest risk zones on these maps. This concentration, combined with the fact that these formerly redlined neighborhoods largely remain economically distressed today, suggest racial bias in the construction of the maps has had important effects over the long run. Using newly digitized data for ten major northern cities, we assess the maps for the importance of this channel in explaining the prevalence of black residents in redlined neighborhoods. We find that racial bias in the construction of the HOLC maps can explain at most a small fraction of the observed concentration of black households in redlined zones. Instead, our results suggest that the majority of black households were redlined because decades of disadvantage and discrimination had already pushed them in to the core of economically distressed neighborhoods prior to the government's direct involvement in mortgage markets. As a result, the HOLC maps are best viewed as providing clear evidence of how decades of unequal treatment effectively limited where black households lived in the 1930s rather than reflecting racial bias in the construction of the maps themselves. We argue that the systemized treatment of neighborhood risk vis-à-vis mortgage lending that was adopted by HOLC and the FHA may have played a central role in locking these patterns of inequality in place.

JEL: R28, N9, J15

[1] Support for this research was provided by the National Science Foundation (SES-1459847). We thank Daniel Aaronson, Shari Eli, Daniel Hartley, Brian Kovak, Jeffrey Lin, Bhash Mazumder, Jonathan Rose, Lowell Taylor, and seminar audiences at Carnegie Mellon, the Chicago Federal Reserve, the Philadelphia Federal Reserve, Yale, George Washington, Pittsburgh, and Arizona for helpful comments. Antonio Diaz-Guy, Jeremy Brown, Andrew O'Rourke, Aly Caito, Loleta Lee, Zach Gozlan, Kaylyn Cameron, Alex Mang, and Sai Konduru provided outstanding research assistance. We are grateful to Prottoy Akbar for assistance with the geocoded address data and to Ken Snowden for giving us the Greensboro, NC FHA map. Corresponding author's email: walshr@pitt.edu (R. Walsh).

I. Introduction

Housing markets have been shaped by racial discrimination and segregation throughout American history. Racially restrictive covenants in deeds, steering by real estate agents, slum clearance and the construction of large public housing projects, the routing of interstate highways, discrimination in mortgage lending, the location of segregated schools, unequal labor market opportunities, threats of outright violence directed at black families, and lack of protection of property and civil rights have all profoundly affected where African Americans lived and the degree to which they could accumulate housing wealth (Kucheva and Sander, 2014; LaVoice, 2020; Brinkman and Lin, 2019; Carruthers and Wanamaker, 2017; Aneja and Avenancio-Leon, 2019). A large literature spanning the social sciences seeks to understand how racial discrimination in housing and different forms of discrimination served to overlap and reinforce each other. In recent years, the discourse has focused on the role the federal government played in the creation of racially segregated neighborhoods across American cities.[2]

Concern regarding the historical role of the federal government has centered on actions taken by the Home Owners' Loan Corporation (HOLC) and Federal Housing Administration (FHA). Established in the 1930s as part of the New Deal's response to Great Depression-related problems in the housing market, these two agencies intervened in U.S. housing markets to a degree unprecedented in American history. The HOLC purchased and refinanced over one million troubled non-farm mortgages and held roughly a tenth of all non-farm U.S. mortgages when it finished lending early in 1936. Established in 1934, FHA-sponsored mortgage insurance grew to cover more than one-third of all new U.S. residential construction by 1949.

[2] Influential books that focus on racial discrimination in the urban housing market related to redlining include Massey and Denton (1993), Satter (2009), Freund (2010), Sugrue (2014), and Rothstein (2017).

To better manage risks associated with their loans, the HOLC subsequently commissioned a series of maps summarizing spatial variation in the riskiness of housing assets in different neighborhoods for over 200 cities. The HOLC maps, in conjunction with contemporaneous maps produced by the FHA, are at the center of current debates regarding the long-run impact of these government agencies on both neighborhood segregation and persistent racial inequality. Black households were concentrated in the highest-risk zones on these maps. This concentration, combined with the fact that these "redlined" neighborhoods largely remain economically distressed today, animates research that seeks to understand the impact of FHA mortgage underwriting behavior on black-white wealth gaps.

Much of this research focuses on the impact of the maps themselves, which have become a visual shorthand for government-sanctioned housing market discrimination in America. The notion that the maps led the FHA to refuse to insure loans in black neighborhoods has spread across academia, the popular press, grass-roots activism, and presidential platforms as part of the debate on the relationship between government policy, racial discrimination and disparate outcomes.[3] Aided by the set of digitized HOLC security zone maps made available by the University of Richmond's Mapping Inequality Project,[4] economists in particular have used the boundaries between security zones to assess the long-run impact of redlining on a range of issues including housing values, neighborhood racial composition, and crime (Aaronson, Hartley, and Mazumder 2019; Anders 2018; Krimmel 2017; Appel and Nickerson 2016).

[3] One influential academic work outside of economics is Rothstein's aptly-named *The Color of Law* (2017). The influential 2014 essay "The Case for Reparations" by Ta-Nehesi Coates prominently cited redlining as rationale for compensating African Americans for the economic harms of twentieth-century discrimination. As a presidential candidate, Elizabeth Warren proposed to compensate residents of formerly redlined neighborhoods in her *American Housing and Economic Mobility Act*. Even the term "redlining" has spread to many related areas; for instance, Zillow and GreatSchools have come under fire for their color-coded school ratings on real estate listings, which critics have referred to as "educational redlining" (for instance, see https://janresseger.wordpress.com/2020/01/13/22699/).

[4] https://dsl.richmond.edu/panorama/redlining/

Implicit in much of the discourse on redlining is the notion that the prevalence of black households in relined neighborhoods stemmed not from the underlying economic disadvantage faced by these individuals and families, but instead reflected discrimination in the construction of the maps themselves. However, to our knowledge there exists no systematic empirical evaluation of how the redlining maps were constructed in the 1930s.[5] Did the majority of black households see their neighborhoods rated as the highest lending risk because they were living there? Or was it the case that low incomes and existing discrimination in employment, education, and housing left black households with few options outside of segregated neighborhoods marked by crowding, dilapidated structures, and depreciating prices that would have been poor lending risks? To what extent were white households with similar incomes and home values to blacks found in redlined neighborhoods? Did the HOLC surveys document existing patterns already in place by 1930 or did the HOLC decisions on how to grade neighborhoods independently disadvantage black households?

To answer these questions, we construct a unique spatial dataset covering individual households and neighborhoods for nine of the ten largest cities in the United States in 1930 and 1940. We also digitized the HOLC maps for these cities along with the extensive information in the survey that accompanied each map. We begin by exploring how housing and economic characteristics varied across security zones, both in levels in 1930 and 1940 and trends across the decade. Using a dataset of almost 300,000 addresses matched to both census years, we then explore how demographic and economic characteristics varied at the boundary of HOLC security

[5] Greer (2012) and Fishback (2014) both estimated regressions based on information from the HOLC surveys, but their results only provide a starting point for the analysis in this paper because the HOLC neighborhood information on incomes, racial shares, and housing values was impressionistic with large amounts of measurement error. Our use of census data on individual households and in the census districts sharply reduces the measurement error problem, as we discuss in Section III.

grades. Finally, we undertake a series of empirical exercises to understand the role that race played in the creation of HOLC's residential security maps. In particular, we assess the relative importance of racial discrimination in the mapping process in explaining the oft-cited fact that the vast majority of black families lived in neighborhoods receiving the lowest (D) rating.

We begin by examining neighborhood-level characteristics. The HOLC gave letter grades A, B, C, and D to neighborhoods with map colors of green, blue, yellow, and red, respectively. In our sample of the largest northern and midwestern cities, over 95 percent of black homeowners lived in the lowest rated "D" zones. However, the vast majority (92 percent) of the total redlined home-owning population was white.[6] Home values, and occupational scores all declined with the security grade, with the "D" zones having the poorest residents and cheapest housing. Thus, black families were already living in the lower-income neighborhoods at the start of the Great Depression. However, we find that redlined white neighborhoods had *better* census economic characteristics on average compared with redlined neighborhoods with an above-average number of black residents, the opposite of what we would expect to see if black neighborhoods had been targeted for the lowest security grade because of race.[7]

We next explore the economic and demographic patterns around the boundaries of HOLC zones. Specifically, we conduct a formal boundary analysis of the differences in economic characteristics of houses very close to C-D borders. Using census data from 1930, years before the maps were drawn, crossing to the lower-graded side of the boundary was associated with a 9 to 10 percent fall in housing prices and a 1.1 to 1.3 point drop in occupational income scores, which implies a drop of roughly $100 in average income in 1950 dollars. Using census data from both 1930 and 1940 to show changes over time, we show that home values were

[6] Throughout the paper we will use the term "redlining" to refer to being assigned to a D security zone.
[7] We use census enumeration districts as our proxy for neighborhood; see Shertzer, Walsh and Logan (2015).

deteriorating, occupational scores were trending downward, and black population shares were increasing on the D-side of the boundary relative to the C-side. Real estate agents thus appear to have accurately delineated neighborhoods on different trajectories that would have impacted lending risk over the long term. We argue that these findings have important implications for the interpretation of "regression-discontinuity" research designs employed in economics research on redlining.

Finally, we explore whether the placement of security zone boundaries reflected racial bias. Using our matched address sample, we explore the location of black households with respect to C-D borders. While in C-rated neighborhoods most black households lived within 200 meters of a C-D boundary, there is little evidence of heaping on the redlined side, suggesting that the maps were not drawn to disproportionately zone black households into the highest risk category. We also perform simple counterfactuals in which we move the boundary between C-D zones to eliminate potentially discriminatory placement. Because black families were typically living deep within neighborhoods that met HOLC's criteria for the highest lending risk, even sizeable boundary shifts would have led to only a small reduction in overall black exposure to redlining. For instance, reassigning blacks who were potentially graded in a discriminatory manner into yellow-lined zones would only have reduced the share of black houses in D zones by about 3 percent.

Our empirical results collectively suggest a more nuanced view of the HOLC and the Federal Housing Administration (FHA), which also made use of neighborhood risk maps.[8] Our findings support the notion that the HOLC maps reflected fundamental and longstanding

[8] The FHA created their own set of maps and manuals for choosing which loans to insure. They had access to the HOLC maps and accumulated additional information from a variety of sources. Unlike the HOLC maps, for which a nearly complete copy was stored in the National Archives, most FHA maps have been lost to history and could not be included in this paper's analyses.

disparities between black and white neighborhoods rather than having targeted black neighborhoods for the lowest security grade. Our contention that the maps exhibit little evidence of explicit racial bias aligns with the conclusions of Hillier's unique (2003a) study of Philadelphia, which found that lenders were avoiding areas that would be redlined before the maps were made. We find little evidence supporting the contention that the individuals who drew the maps made decisions about rating neighborhoods that would have made black access to FHA lending worse in 1937 than it had been in 1930.

A more complete understanding of the emergence of redlining should build on the private market forces that generated disparities between black and white neighborhoods before the Great Depression, and the persistence of such disparities. The FHA made little effort to ameliorate the impacts of white avoidance of black neighborhoods, which had dramatic adverse impacts on housing prices in these areas (Akbar et al. 2020). The FHA lending guidelines instead emphasized avoiding lending in neighborhoods with uncertain price trajectories, which were often black neighborhoods in part because of actions taken by the federal government elsewhere. In particular, demolishing black neighborhoods to make space for highways, concentrating urban poverty through public housing policy, and the acceptance of labor market discrimination and enormous inequality in education at every level continued to disadvantage black families and sharply limit their ability to accumulate housing wealth. The federal government did not move to seriously address these disparities until the 1960s, by which point racial segregation had become deeply entrenched in American cities.

II. Background and Historical Context

a. Home Finance Before the Crash

By the 1920s, academics and other real estate professionals understood neighborhoods to be dynamic and expected housing values to change over time (Babcock 1932, Burgess 1928, Cressey 1938, Gibbard 1941, Schietinger 1951 and 1954). In their view a city was a composition of zones, and property values within those zones were affected by the succession process: the introduction of new demographic groups, the reaction or abandonment from the original residents, and the achievement of a new equilibrium of communal stability. Most neighborhood successions were gradual, but Frederick Babcock (1932), later the FHA's head of underwriting, argued that one factor that hastened neighborhood transition was race (see also Burgess 1928). In *The Value of Real Estate* Babcock writes, "Most of the variations and differences between people are slight and value declines are, as a result, gradual. But there is one difference in people, namely race, which can result in a very rapid decline" (1932, p. 91).

The dynamics of neighborhood successions meant that lenders considered long-term home loans to be more risky than short-term loans, and that some of this uncertainty was driven by the potential for racial transitions that would impact property values. Through the 1920s, lenders in the mortgage market had mitigated the risk of neighborhood succession by focusing predominately on localized lending and short-term loans. At the time, a common mortgage was a 3- to 5-year straight, "balloon" loan through an insurance company or a state bank (Snowden, 2013). On a 5-year $1,000 loan the borrower paid back only interest for five years and then paid the principal of $1000. Often, these loans were rolled over with some adjustments if the borrower paid part of the principal.

Building and loans (B&Ls) were the leading institutional mortgage lenders of the day, accounting for about 40 percent of new mortgage loans in the 1920s. They focused on local markets and provided longer-term mortgages that combined a balloon loan with a "share

accumulation contract," which required investments in that B&L's shares. For a $1,000 loan the borrower each month paid the interest plus bought a share in the B&L. These payments continued until the value of the shares plus dividends paid on the shares equaled $1,000 and the loan was finished. The full amount of the principal was owed throughout the loan until the $1,000 was paid. Depending on the dividend rate, the mortgage lasted about 11 to 13 years.

The B&Ls experimented with the current-day loan form, the direct reduction loan (DRL), as early as the 1880s, but most did not use them until the 1930s (Rose and Snowden 2013).[9] A direct reduction loan required the borrower to make equal payments throughout the life of the loan. Each payment was composed of interest and a payment that reduced the amount of principal owed, hence the term "direct reduction." Over time the part of each payment going to reducing the principal rose and the interest part declined. Direct reduction loans became common when they became the primary lending vehicle used by savings and loans, the institutions that generally outcompeted B&Ls in the 1930s. The HOLC used them in the 1930s as well, and the FHA would only insure such "conventional" direct reduction loans.

During the 1920s, the majority of mortgage loans were thus made by local institutions – including building and loans, mutual savings banks, and state banks – with deep knowledge of local housing markets.[10] B&Ls held more mortgage debt on single-family homes than all other financial institutions combined, with approximately 40 percent of home buyers borrowing through a B&L for some part of their mortgage (Snowden, 2010). B&Ls typically started out as small organizations created by surveyors, title specialists, attorneys, real estate and insurance agents, homebuilders, and building material suppliers who organized and operated B&Ls on a

[9] While less widespread, private lenders had also begun to adopt long-term, locally issued loans in the 1920s (Van Dyke, 1929).
[10] National banks could not loan on real estate until 1927.

part-time or voluntary basis. The borrowers were the owners and members, and they met regularly to pay weekly or monthly dues. Often, the members were part of the same social or work network. The networks and the focus on local housing markets allowed the B&Ls to effectively manage the investment risk associated with changing neighborhood trends and demographics while simultaneously offering long-term loans (Snowden 2003).

b. The Great Depression and the HOLC

During the Great Depression, real GDP fell by 30 percent and unemployment rates rose above 20 percent. At the same time roughly 40 percent of home borrowers fell behind on their mortgage payments. Families struggled to refinance their loans and B&Ls became "frozen," unable to make new loans while servicing existing loans (Rose and Snowden, 2013). The federal government responded by establishing the Homeowners' Loan Corporation (HOLC) as part of the Federal Home Loan Bank Board (FHLBB). The HOLC bought over a million loans from lenders, replacing toxic assets on their books. The HOLC then refinanced the loans for borrowers using 15-year direct reduction loans with five percent interest rates, terms that were better than in the regular market (Fishback, Rose and Snowden 2013).

In contrast to the typical housing lenders, the officials at the HOLC needed to understand the risk levels in hundreds of idiosyncratic housing markets throughout the country. How could they obtain all the local information needed to evaluate the loans, the properties, and the riskiness of the borrowers? Their solution was to set up over 200 local offices and hire over 20,000 people throughout the country to work with local appraisers while negotiating with lenders and borrowers. The HOLC finished the refinancing project early in 1936 and then shifted to servicing the loans.

By the middle of 1935, refinanced borrowers were already having problems repaying their loans. In September 1935, HOLC officials undertook a City Survey program to collect information on local real estate and mortgage conditions "to successfully establish policies with respect to the collection on HOLC loans, the management and ultimate sale of acquired real estate as well as to the rehabilitation of the savings and loan industry…" (National Archives Undated, 1). Over the next five years they conducted the surveys and produced general reports about the economies of over 200 cities, surveys of all mortgage lenders about their lending and assets, and detailed descriptions of neighborhoods that they used to develop risk grades for lending in each neighborhood in the city.

To visualize the geography of the risk grades, the HOLC officials developed "residential security maps" that assigned colors to each grade: green (A); blue (B); yellow (C); and red (D). The green (A) rating signified the lowest level of lending risk, while a red (D) rating signified the highest risk. To aid in the development of the maps, the City Survey Program consulted with local real estate professionals including local bank loan officers, city officials, and realtors in every major city to assess perceived lending risk on a neighborhood-by-neighborhood basis. Neighborhoods were categorized based on several criteria including the age and condition of housing, access to amenities such as transportation and parks, the neighborhood's racial and ethnic composition, and the economic status of neighborhood residents. Field agents also documented the availability of mortgage funds within each of the security grade zone.

Many of the printed forms for the neighborhood information offered the following descriptions on the back. The A areas were "'hot spots'; they are not yet fully built up. In nearly all instances they are the new well planned sections of the city, and almost synonymous with the areas where good mortgage lenders with available funds are willing to make their maximum

loans to be amortized over a 10 to 15 year period, perhaps up to 75-80 percent of the appraisal." The B areas were "completely developed, like a 1935 automobile – still good, but not what the people are buying today who can afford a new one…good mortgage lenders will have a tendency to hold loan commitments 10 to 15 percent under the limit." The C areas were characterized by "age, obsolescence, and change of style; expiring restrictions or lack of them; infiltration of a lower grade population; the presence of influences which increase sales resistance, such as inadequate transportation, insufficient utilities, perhaps heavy tax burdens; poor maintenance of homes, etc." Lenders were more conservative and held loan commitments under the lending ratio for both A and B zones. Lastly, D "areas represent those neighborhoods in which the things that are now taking place in the C neighborhoods, have already happened…The areas are broader than the so-called slum districts." Loans were made on the most conservative terms and some lenders refused to make any loans in D zones.[11]

We emphasize that these maps were created after the HOLC had finished making loans and did not guide the institution's refinancing project.[12] In fact, the HOLC itself made substantial loans in neighborhoods that would later come to be D-rated. For instance, Amy Hillier's (2003b) study of Philadelphia found that HOLC made 60 percent of its loans in future D-rated areas. Working with a sample of loans matched to census records, she also showed that

[11] From the Report on Decatur, Illinois by the Division of Research and Statistics with cooperation of the Appraisal Department, July 25, 1937. HOLC City Survey Records, Record Group 39. These statements were similar to material found in archival material and journal articles from the 1930s, which suggests that lenders were avoiding redlined areas before HOLC made its maps (Hillier, 2003a).

[12] Confusion on the history of the HOLC appears frequently in the discourse on redlining. For instance, a recent article in Bloomberg's CityLab stated that "In 1933, the federal Homeowners Loan Corporation began playing a huge role in directing mortgage loans to mostly white families in danger of foreclosure. The HOLC's real estate survey maps codified the redlining practices of the private real estate appraisal industry for determining where it would focus its investments" (https://www.bloomberg.com/news/articles/2017-02-21/mapping-louisville-s-redlining-history), suggesting that the maps guided HOLC lending decisions. Similarly, an essay in the *New York Times* "1619 Project" claims that the HOLC excluded black neighborhoods from government-insured loans after shading them with red on a map: (https://www.nytimes.com/interactive/2019/08/14/magazine/racial-wealth-gap.html). The HOLC did not insure mortgages.

blacks and immigrants were overrepresented in the pool of homeowners who received refinancing from the HOLC. The black share of HOLC loans was higher than the black share of homeowners in 47 cities and lower in only 17 cities studied by Michney and Winling (2019, pp. 10-11). Nationwide, black households accounted for 4.5 percent of mortgages held by the HOLC in 1940, compared with only 2.6 percent held by all other lenders (U.S. Bureau of the Census, 1943, pp. 7, 9). The 4.5 percent matched the black share of nonfarm homeowners in 1930 and in 1940. The HOLC's lending patterns thus demonstrate a substantial degree of assistance to black mortgage holders.

c. The FHA

In 1934 the Federal Housing Administration was created, also under the FHLBB, and it began insuring mortgage loans for home maintenance and rehabilitation. In 1935, the FHA began insuring home mortgages, mostly for new construction. The FHA was originally highly risk averse and avoided insuring risky properties in order to keep foreclosure rates down. In the FHA's *Fifth Annual Report* FHA Administrator Steward McDonald described the purpose of the FHA:

> "Through a policy of insuring only such mortgages as meet reasonable tests of economic soundness, and of rejecting all others, home ownership is aided on a basis where the borrower assumes an obligation within his reasonable capacity to pay, and where the home that purchases meets high standards of construction and neighborhood. The Administration's efforts in this direction, especially in insisting on competent lay-out and reasonable neighborhood protection for new subdivisions have been more effective than ever before, because of the large increase in the number of new homes financed with insured mortgages." Federal Housing Administration (1939, pp. vii-viii).
>
> "In each of the 48 States the Federal Housing Administration has established standards to eliminate jerry building. Through its land planning activities, the Administration is assisting in setting up effective barriers against neighborhood blight and the development of future slums. These activities, carried on for the purpose of protecting the mutual interests of the borrowers, the lenders, and the

> mortgage-insurance fund, are largely responsible for the insistent and constantly increasing demand for the services of the Federal Housing Administration." Federal Housing Administration." Federal Housing Administration (1939, pp. vii-viii).

When the FHA was established in 1934, officials recognized that they had to learn a great deal about local markets when insuring loans; therefore, they embarked on their own studies of local markets. The question of how much the HOLC and the FHA cooperated on their neighborhood studies is still debated by scholars. The historian Kenneth T. Jackson (1985, p. 203) suggested that "the HOLC appraisal methods, and probably the maps themselves," were adopted by the FHA. However, the HOLC itself considered a large amount of the information collected, including the maps, to be confidential, and the maps were almost entirely held within the HOLC, with some copies available to other parts of the FHLBB (Hillier, 2003a; National Archives undated). The finding aid at the National Archives (undated) describes in detail who had each copy and how the copies were collected and destroyed until they were declassified sometime before 1985. Nearly all of the copies of the confidential information were held by HOLC officials. FHA officials were given one copy of the HOLC full reports, which is the only material with information on race and ethnicity of the residents, and three copies of the maps without area descriptions.[13]

In determining their own views of neighborhoods when appraising individual loans, the FHA Division of Economics and Statistics combined the HOLC information with their own local field studies and surveys of real property that they designed and asked the Works Progress Administration (WPA) to perform for 74 cities in 1938.[14] The FHA produced their own set of neighborhood risk maps, but very few of these maps have been found by scholars. Two

[13]A large number of D neighborhoods contained only a few blacks, so the area descriptions were the only way for someone to determine which areas were given D grades because of the presence of blacks.
[14] Federal Housing Administration (1939, 40-42). See also Hillier (2003a) and National Archives (undated)

exceptions are the maps for Greensboro, NC and Chicago, IL which we reproduce in Appendix III alongside the HOLC security zone maps for these cities. The maps do bear a broad resemblance in terms of where risk is concentrated. However, there are also significant differences in the location of boundaries between zones.

Economists wishing to study the long-term impacts of redlining have relied on the HOLC security grade maps in their analyses because scholars have been unable to find the FHA neighborhood risk maps that were actually used to guide mortgage insurance decisions. The many disagreements in precise boundaries between the HOLC and FHA maps for Chicago and Greensboro thus raise questions about the interpretation of typical "regression-discontinuity" analyses performed on the HOLC security grade boundaries. If the HOLC maps accurately delineated neighborhoods on different trajectories by 1930, then the persistence of such local dynamics could generate persistent disparities in economic outcomes even over short distances. We explore this question in more detail below.

III. Data

To better understand the HOLC's risk assessment maps, we construct a novel spatial dataset linking data from the 1930 and 1940 censuses to HOLC residential security zones for a sample of major northern cities that have been the focus of concerns related to redlining. In earlier work we digitized the underlying census enumeration districts for this sample of major cities in the North (Shertzer, Walsh, and Logan, 2015).[15] Our sample for this paper covers Baltimore, Boston, Brooklyn, Chicago, Cleveland, Detroit, Manhattan, Philadelphia, Pittsburgh,

[15] The census enumeration districts were small administrative units used by the census for enumeration purposes. They typically covered 1500 people in urban areas. We also digitized the enumeration districts for Cincinnati. However, the HOLC map for Cincinnati seems to have been lost to history and could not be included in the sample for this paper.

and St. Louis. These were the largest nine northern cities in 1930, covering nearly 18 million people (about half of the total in the largest 100 cities) and about 38 percent of the urban black population living outside of the states of the former Confederacy.

The demographic data used for the paper have three components. The first is census data aggregated at the enumeration district level. Enumeration districts were small administrative units used by the census and typically covered one to four city blocks in urban areas. We use them as our proxy for neighborhood. The second piece is at the level of the HOLC security zones. To study these areas, we use data both from the HOLC surveys themselves and from census data on individual households that we aggregate to these zones, both of which are discussed in more detail below. The third component is a dataset containing census data for individual addresses that were matched across the 1930 and 1940 censuses.

We digitized both the residential security maps and the detailed survey that accompanied the maps for each of our cities, yielding observations for a total of 927 HOLC security zones.[16] The date of creation for the maps themselves ranges from 1937 to 1940. The associated surveys documented housing characteristics, including housing prices, construction type (brick, frame, or other), and the general state of repair (excellent, good, fair or poor), as well as population characteristics, including rough estimates of the typical occupation, average income, and racial composition of neighborhood residents. The surveys include retrospective data reaching as far back as 1929 for the highest and lowest housing value for up to three types of housing and the highest and lowest rents for up to three types of rental properties. These surveys also report the perceived future desirability trend for each zone.

[16] Since we began this project, the University of Richmond digitized a larger sample of HOLC security zone maps. These can be found at https://dsl.richmond.edu/holc/

Descriptive characteristics from HOLC surveys are presented by security grade in the first 4 columns of Panel A in Table 1.[17] The mid-points between the highest and lowest house values and rents are for the year closest to 1935 in each survey. Recall that neighborhoods with an "A" rating signified the lowest level of perceived lending risk, while a "D" rating signified the highest. These summary statistics based on HOLC survey data show that income, housing values, and rents were all negatively correlated with perceived lending risk, while the shares of black and foreign-born residents were associated with higher risk, with black households located almost exclusively in the highest risk "D" neighborhoods. Only ten zone C (yellow) neighborhoods out of a total of 286 were reported to have a black population share above 4.3 percent (the mean across all HOLC zones). The correlations we document here are consistent with earlier analyses of race and security grade determination (for instance Greer, 2012).

To augment the neighborhood-level information reported in the HOLC surveys, we attach HOLC security zone identifiers to individual census observations by overlaying the HOLC maps on census enumeration district (1930) and census tract (1940) maps for our nine-city sample. We proceed by calculating the share of each enumeration district or tract that lies within each security grade and then attach individuals from the census to each security grade using areal interpolation. Summary statistics from the 1930 census data aggregated to security grade are presented in Columns 1-4 of Panel B in Table 1.

The census data confirm that, as of 1930, the majority of black households lived on city blocks that were destined to be shaded in red. Over 97 percent of black individuals and 95 percent of black-owned homes ended up in red-shaded HOLC zones. While blacks were highly

[17] The Appendix contains a more detailed description of the data construction. Summary statistics for additional variables are presented in Appendix Table A1. Additional variables include construction type (brick, frame, other) and the general state of repair. Green (A) and blue (B) zones were more likely to have brick houses in good condition, while yellow (C) and red (D) zones were more likely to consist of frame houses in fair or poor condition.

concentrated in red-shaded neighborhoods, it was still the case that the majority of individuals in red-shaded areas were white. Fully 49 percent of the 9 million white individuals and 39 percent of the 700,000 white homeowners in our 1930 sample were in neighborhoods that were shaded red in the maps. Because of their much larger numbers, these whites accounted for 82 percent of individuals and 92 percent of the owned homes in the D red-shaded areas.

The additional census neighborhood variables summarized in Panel B are consistent with the survey data reported in Panel A. Median housing values, median rents, occupational income scores, and the share of owner-occupied housing are negatively correlated with low security grades, while the share of black and foreign-born residents is positively associated. To further understand the role of race in shaping the HOLC's redlining maps, columns (5)-(8) of Table 1 separately show summary statistics for the two highest-risk security zones (C and D) while splitting them into groups of neighborhoods that had percent black shares above or below the average percent black at the zone level of 4.3 percent. A pattern emerges in these summary statistics. In particular, red-shaded zones with a high black share appear to be more economically disadvantaged than red-shaded zones with a low black share, a result that holds in both the HOLC surveys *and* the census data.

To look more closely at the factors influencing the assignment of households around the boundaries between zones, we use a dataset of geocoded addresses that were matched between the 1930 and 1940 census waves based on their addresses by Akbar et al. (2019). Because each observation in the panel was geocoded, each address is correctly assigned to the HOLC security grade and its closest HOLC zone boundary, and the distance to the boundary is computed.[18]

[18] We drop all houses that are within 30 meters from an HOLC boundary to mitigate any concern over measurement error and to prevent any comparisons of households directly across the street from one another. We choose 30 meters since this is the average depth of a household plot.

The summary statistics by security grade for the address sample, which are reported in Appendix Table A2, show overall patterns similar to the summary statistics reported in Table 1. The security grade classifications are associated with both economically and statistically significant differences in observable characteristics, with poorer neighborhoods rated a worse lending risk.[19]

IV. Analysis

Our empirical analysis seeks to understand why so many black households were in the red districts under the HOLC maps. We focus on two competing, but not mutually exclusive, hypotheses. First, HOLC assessors may have assigned lower risk neighborhoods with large numbers of black households to the highest risk grade because they were racially biased, either explicitly due to animus or implicitly because of racially-driven beliefs about the quality and prospects of black neighborhoods. Alternatively, the economic hardships imposed on black families by decades of discrimination in education, protection of property rights, and employment combined with racial barriers in housing markets could have left them with few options outside neighborhoods that met the criteria for the highest risk rating, independent of racial composition.

a. Survey Quality and Boundary Determination

We begin with a discussion of how the neighborhood surveys undertaken by the HOLC compare to security zone-level variables we constructed using census data from 1930 and 1940. We then consider the salience of the boundaries chosen by the HOLC for these zones. While we believe it is unlikely that the production of the HOLC maps between 1937 and 1940 could have

[19] For example, housing values averaged around $9400 in 1930 for zones coded as green (A) while only averaging $5400 in zones coded as red (D). The percentage change of each variable is also reported. The impact of the Great Depression is clear in our data. Average housing values decreased by 20% in green zones between 1930 and 1940 but decreased by 43% in red zones.

had any meaningful impact on neighborhood demographics reported in the 1940 census, we highlight the 1930 census data to be as conservative as possible regarding issues of reverse causality.[20]

To get a better sense of the relationship between data reported in the potentially subjective HOLC surveys and neighborhood conditions from the 1930 census, Appendix Figures A3 through A6 compare the cross-neighborhood empirical distributions for median housing values, median rents, share black and share home-owners from the census with those reported on the HOLC surveys for each of the 4 security zone classes. For the HOLC survey data, we proxy for median values and rents using the midpoint between the reported highest and lowest housing values.[21] Housing value and rent distributions track reasonably well across the two data sources, with the survey-based distributions' leftward shift to be expected given the secular decline in housing prices that occurred between the 1930 census and HOLC survey years later in the decade (Fishback, Rose, and Snowden 2013; Fishback and Kollmann 2014). The distribution of share black also tracks well across the two measures despite there being relatively less within-security grade variation in both the census and HOLC data. The distributions of homeownership vary quite substantially, likely because home ownership rates dropped substantially after 1930 and the HOLC consultants focused on single-family homes when filling out the survey (Fishback, Rose, and Snowden 2013).

[20] We highlight here that due to the retrospective nature of much of the HOLC survey data, the data summarized in Panel A of Table 1 are generally from the middle of the decade, meaning that the 1930 census data are roughly as temporally proximate to the data we summarize from the HOLC surveys as are the 1940 census data. Thus, absent potential concerns about reverse causality, 1930 and 1940 were equally appropriate sources of comparison, making the choice to focus on 1930 straightforward.

[21] There are several reasons to expect that the HOLC surveys and census information should not match perfectly. The census means and percentages are calculated from surveys of individual households, while the HOLC data are the rough estimates for the neighborhood made by a group of housing market professionals. Further, for values and rents, the medians from the census and the midpoints from the HOLC surveys are different measures of rents and housing values. In addition, there is some slippage because of the interpolation from the census districts to match the HOLC neighborhood boundaries.

In addition to these static measures, the HOLC's zone assignment decision was also a function of the reported neighborhood desirability trend over the next ten to fifteen years. For example, 94 percent of neighborhoods assessed as having downward trends were classified into security zones C and D, while 86 percent of neighborhoods identified as having upward trends were classified into zones A and B, with none of the latter being classified into zone D. To investigate the relationship between HOLC's trend classification and actual neighborhood change across the 1930s, we use the 1930 and 1940 census data interpolated to HOLC zones to estimate the following linear regression:

$$y_{ic1940} = \alpha + \sum_{j} \beta_j I_{trend_i=j} + y_{ic1930} + \gamma_c + \epsilon_{ic} \quad (1)$$

where y_{ic1940} is a census outcome for zone *i* in city *c* in 1940 and y_{ic1930} is the lagged value of the census outcome variable, $I_{trend_i=j}$ is an indicator that the HOLC survey placed neighborhood *i* in trend category *j*, and γ_c are city fixed effects which control for any unobservable characteristics that are constant across all security zones in a given city. The coefficients of interest are the $\beta_j's$, which identify the relationship between the different predicted trends and the ten-year change in the dependent variable.

Figure 1 summarizes the results of these regressions for log median housing price, log median rent, share black and occupational income scores (a commonly used proxy for income, based on an individual's occupation).[22] These results suggest that the predicted trends captured statistically significant and economically meaningful differences in neighborhood trajectories. Each increase in the optimism of HOLC predictions is associated with a rise in the growth rates in housing values, rents, and occupation codes. The gap between HOLC predictions for upward

[22] See Saavedra and Twinam (2020) for a discussion of occupational income scores. It is an estimate of the median income for a given occupation in 1950, measured in 100's of 1950 dollars.

trend and downward trend was associated with a growth rate that was 46 percent higher for housing prices, and 22 percent higher for rents, as well as a rise in the occupational score that was the equivalent of an additional $400 (in 1950 dollars). If increases in share black were associated with the HOLC predictions of the trend in future desirability, we would have expected the values in Panel D of Figure 1 to be below 0 and to become more negative as the predictions moved upward. Instead, the coefficients for "slightly downward" and "static" were positive, and none of the coefficients were statistically different from zero.[23]

We next explore the placement of the HOLC security zone boundaries. For parsimony, we focus our boundary analysis on neighborhoods that were assigned to security grades C and D since these zones capture nearly all the variation in racial composition across security grades. We begin by using our matched sample of geocoded single-family homes to explore variation in house and household characteristics across these boundaries. Figure 2 plots means and 95-percent confidence intervals from 50-meter bins of log house prices, log rents, occupation score, and percent black as a function of distance from each side of the C-D boundary using the address-level data for 1930. The bin means clearly show substantially higher home values, rents, and occupational scores and substantially lower shares of blacks on the C-side of the boundary than on the D-side. Seven to 10 years before the HOLC surveys and maps were developed, stark differences in these features were already in place. The situation using 1940 Census data are similar and are reported in Appendix Figure A8.[24] We provide a trend analog to Figure 2 in

[23]The results of an additional analysis that uses security grade fixed effects in the regression to focus on changes over time within security grade show the same patterns in Appendix Figure A7. The gap between HOLC predictions for upward trend and downward trend was associated with a growth rate that was 31 percent higher for housing prices, and 10 percent higher for rents, as well as a rise in the occupational score that was the equivalent of an additional $378 (in 1950 dollars). The coefficients for share black were very similar to the ones in Figure 1. The regression results underlying Figure 1 and Appendix Figure 7 are reported in Appendix Table 3.

[24] The boundaries themselves were drawn at most three years prior to the 1940 census making the 1940 data, all else equal, a better measure for the boundary analysis. Greer (2012) documents that existing homes were largely ineligible for FHA lending (which could have been influenced by the HOLC maps), which we have confirmed with

Figure 3 where we report 1930-1940 changes. While there appear to be some differential trends between the C and D zones, particularly for race of occupant, the raw data plotted here do not reveal any stark jumps in trends at the C-D boundary.

To provide a more statistically grounded assessment of boundary discontinuities, Table 2 presents results from a standard regression discontinuity (R-D) model of the following form:

$$y_{ij} = \alpha + \beta lgs_{ij} + \rho dist_{ij} + \varphi dist_{ij} * lgs_{ij} + \gamma_j + \epsilon_{ij}, \quad (2)$$

where y_{ic} is the outcome for address i near boundary j, and $dist_{ij}$ is a measure of the distance of address i to boundary j. lgs_{ij} equals 1 if address i is on the lower-grade (D) side of boundary j, γ_c are boundary fixed effects, and ϵ_{ij} is the error term. The coefficient of interest in equation (2) is β, which measures the extent to which addresses on the lower-grade side of a boundary were discretely different from addresses on the higher-grade side. The results use the optimal bandwidth selection proposed by Calonico, Cattaneo, and Titiunik (2014). The results are presented for both 1930 data and 1940 data, in addition to a trends model which replicates the 1940 iteration of equation (2) but includes as a control the 1930 value of the dependent variable (thus replicating the approach taken in equation 1). We document the robustness of all results in this table to a wide range of bandwidth choices in Appendix Figures 9 through 11.

The R-D results reported in Table 2 reinforce the visual evidence from Figure 2 and further suggest trend discontinuities that are not immediately apparent in Figure 3. In particular, the boundaries drawn by the HOLC captured statistically significant and economically meaningful discrete changes in important neighborhood characteristics. We focus on the

our own analysis of FHA reports. We thus believe our sample was largely excluded from FHA activity prior to 1940, and therefore that the maps themselves could have had little causal impact on the evolution of neighborhood characteristics over the 3 short years between 1937 and 1940. As a result, the 1940 data should more accurately reflect pre-existing conditions that motivated delineation of the security grade zones, compared with the 1930 data. We nonetheless note that the estimates from the two samples are overall quite similar.

estimates from specifications (2) and (4), which control for boundary fixed effects. The results for specifications (1) and (3) without the fixed effects are similar. Setting aside race for a moment, crossing to the lower-graded side was associated with a 9 or 10 percent fall in housing prices and a 1.1 to 1.3 point drop in occupational income scores, which implies a drop of roughly $100 in average income in 1950 dollars. Crossing the boundary into the D zone meant that rents fell 1.6 to 3.7 percent (but only the result for 1940 values is statistically significant). The results for the trends from 1930 to 1940 are qualitatively similar but with smaller magnitudes and a statistically insignificant negative effect for rents

The is also clear evidence of discontinuities in racial composition at the border, with a 4.9 percentage point jump in percent black in 1930 and a 3.5 percentage point jump in 1940 on the lower-graded side. Similarly, the trends analysis finds a discrete 1.6 percentage point jump in the rate of increase in percent black. As a point of comparison with the economic variables, the racial discontinuity in 1930 (1940) is 14.1 percent (11.5 percent) of the standard deviation across all locations in the C and D zones. The corresponding percentage for prices is 14.1 percent (14.6 percent), for rents is 2.85 percent (6.9 percent) and for occupation scores is 13.2 percent (11.5 percent).

Taken together, these results highlight the difficulty of separating the relative roles of race and economic distress in the generation of the HOLC maps. The evidence suggests that the chosen boundaries identified points of abrupt neighborhood transition, both economically and racially. We explore the racial dimensions of HOLC mapping in more detail in the next section, focusing on the relationship between security grades and the overall economic disadvantage facing black families at the time of the maps' creation.[25]

[25] One possible concern regarding the border analysis is that these results could be driven by extremely "thick" borders, meaning borders that are associated with significant physical barriers (i.e. railroad tracks, major roads and

b. Factors Related to the Concentration of Black Families in D Zones

Our further analysis of the racial dynamics of HOLC map creation proceeds along two dimensions. First, we search for evidence of the importance of race at the neighborhood level by comparing the relative non-racial character of high and low black locations that were differentially classified into security grades C and D. Next, taking the overall neighborhood-level grade assignments as given, we ask: what portion of the concentration of black families in redlined neighborhoods can be explained by the racially motivated assignment of these specific boundary locations? If the HOLC local experts had used neighborhood racial composition as a factor in determining risk grades, we would expect to find on average that redlined neighborhoods containing large numbers of black households with higher incomes and more valuable homes than their white counterparts. This pattern would arise because the use of race as a factor in redlining would implicitly mean that some neighborhoods with more black households would have been assigned to D security grades even though their underlying economic situations were more in line with those of a C-security-grade neighborhoods.

For this analysis, we divide our cities into small neighborhoods, each of which will constitute a single observation. We use the 100 percent count 1930 census data aggregated to the enumeration district-level for this analysis.[26] We first develop an index of economic distress for each enumeration district based solely on non-racial data. We begin this process by estimating the following linear probability model.

---

rivers). Our results are robust to the exclusion of such border segments. See: Appendix Table 4 and Appendix Figures 12-14.

[26] This analysis is done using 1930 enumeration district level data because 1940 aggregated census data is only available at the tract level. Given the larger spatial area covered by tracts, the 1940 data was less-well suited for this analysis. We replicate this analysis using 1940 census tracts. These results are presented in Panel B of Appendix Figure A15. In both cases we restrict the sample to enumeration districts (1930) or tracts (1940) that had at least 95 percent of their territory uniquely fall into a C or D security grade. For our boundary analysis, we use geocoded addresses so there is no difference in the spatial suitability of the 1930 and 1940 data.

$$I_{ED_i \in D} = \beta' X_i + \varepsilon_i, \quad (3)$$

where $I_{ED_i \in D}$ is an indicator variable with a value of one when an enumeration district is assigned to a D-grade security zone and zero when assigned to a C-grade zone. In this regression, X_i is a vector containing the following non-race enumeration-district characteristics: share foreign born, share homeowners, average age, average occupation score, average rent, average sales price, and labor force participation rate. We then use the estimated coefficients from this linear probability model to predict the probability that each enumeration district in the sample would be classified into a D security grade, essentially yielding an index of perceived economic distress that doesn't explicitly include race as a factor.[27]

The goal is to compare the distribution of this enumeration district-level index (predicted probability of being ranked D) across areas with large and small black populations. To that end, after computing the estimated probabilities, we divided the sample of enumeration districts (EDs) into two groups: EDs with a greater than 15 percent black population share and EDs with black population shares of 15 percent or less (15 percent is the average black share in D zones, but results are robust to a wide range of racial cutoffs).[28]

In Panel A of Figure 4 the sample contains EDs that were assigned a D rating by the HOLC. The continuous line shows the distribution of our predicted index for EDs with black shares of more than 15 percent, while the broken line shows the distribution of the index for EDs with black shares of 15 percent or less. The unbroken line for a higher black share is

[27] We note that one should be careful interpreting the specific coefficient estimates from this model as the exclusion of race will lead to omitted variable bias. However, in our case, we are only using these coefficients as weights for constructing an index of neighborhood characteristics associated with economic distress. We further note that all coefficient estimates have the expected sign.

[28] These results are robust to varying both the 95 percent coverage criteria for inclusion in the sample and to varying the 15 percent racial threshold (see Appendix Figure A15). Another potential concern is that the results may be driven by systematic differences across cities in racial composition and overall economic distress. In Appendix Figure 16 we show that these results also hold when city fixed effects are used to net out city-level factors.

concentrated to the right of the broken line for the lower black share, which shows that the homeowners in EDs with higher black population shares typically had worse economic characteristics, the opposite of what we would expect to see if black neighborhoods had been disproportionately targeted for the D rating. The comparison of groups with high and low black shares for EDS with actual C ratings in Panel B of Figure 4 shows similar results. The same pattern holds using 1940 census tract data as opposed to 1930 ED data (see Figure A16).

These results once again show that black households were already concentrated in the most economically challenged neighborhoods in these cities seven to ten years prior to the development of the HOLC maps. Thus, the assignment of a D rating and red shading for high risk to those neighborhoods where the share of black families was higher would almost certainly have happened even if the HOLC decision makers had not known the race of the families in the neighborhood.

This finding is reinforced in Figure 5, which plots the share of black-occupied single-family homes in groups at various distances from the CD boundary. As the plot moves from the vertical line at 0 to the right, the distance from the C-D boundary rises. As the location moves further into the district, the share of black families increases, with the share of black families being highest on the far right of the graph in the physical heart of the largest contiguous D-rated areas. It is precisely the high concentration of black families in the core of these economically distressed neighborhoods that underpins the findings in Figure 4 that on average, D neighborhoods with high black shares had lower incomes and housing values than the D neighborhoods with low black shares.

c. The Location of Boundaries

We now turn to an evaluation of the choice of specific boundary locations. Given the results of our analysis of the security grade assignment, it is likely that for most black families their homes were assigned D-ratings largely because they were embedded in the center of neighborhoods with the lowest (non-racial) socio-economic characteristics. Such a finding, however, does not preclude the possibility that the HOLC shifted its zone boundaries in one direction or another so that more black households would be put into D-rated neighborhoods. In fact, a close examination of Figure 5 potentially provides evidence of such behavior with some apparent bunching of black households just inside (within 100 meters) of the redlined side of the border.

The coefficients in the R-D boundary analysis presented in Table 2 identify the existence of discrete changes in home value, rents, income and race at the specific boundaries chosen by the HOLC to demarcate between C and D zones. We begin here by assessing how the coefficients in Table 2 using the actual boundaries selected by the HOLC compare with coefficients from R-D analysis when the location of the C-D boundaries are randomly assigned. Specifically, we produce a set of hypothetical boundaries, uniformly distributed at 5-meter intervals within a 250-meter buffer of the true boundaries. This procedure yields a set of 101 distinct possible boundaries between each adjacent C and D zone. We then randomly choose one boundary from this set for each zone pair and re-estimate the models from Columns 2 and 4 of Table 2 on this set of hypothetical boundaries. Replicating this process 1,000 times provides an empirical estimate of the distribution of coefficients for the economic and demographic discontinuities that arises when the C-D boundary location is randomly assigned that many times. These distributions (along with the estimated discontinuity at the actual boundaries) are presented in Figure 6 for both 1930 and 1940. We also use this distribution to compute a one-

tailed empirical p-value (in parentheses below each panel) that measures the probability that random assignment would have led to the boundary that the HOLC actually chose and therefore produced the coefficient from Table 2.

The empirical distributions for all four measures (value, rent, income and race) are approximately normal and roughly centered at zero. Further, the estimated p-values for the actual boundaries suggest that the actual HOLC boundary locations were not randomly assigned because the p values for the dimensions of housing value (p-value .016 for both year), income (p-value .005 and .011) and race (p-value .011 and .060) are low. HOLC agents specifically chose to place boundaries at locations where discrete changes occurred along economic and racial dimensions. Focusing on race, the evidence presented in Figures 5 and 6 suggests that race itself played a role in determining the specific locations of some security grade boundaries. Visual inspection of Figure 5 suggests that this process led to a bump upward by roughly 5 percentage points in the probability that a household was black located between 40 and 100 meters inside the D zone side of the C-D boundary.

Motivated by this clumping of black households within 100 meters of the boundary, we conclude our analysis by attempting to gauge the importance of race-based boundary selection through the construction of two counterfactuals. First, we consider the outcome if *all* C-D boundaries had been shifted 100 meters into the D-zone side of the boundary. This exercise ignores all other information about occupations and housing values in these locations and shifts the bump in black households from the D zone to the C zone. In doing so, we shift 2,474 black households (19.8 percent of all D-zoned black households) and 868 black-owned homes (20.9 percent of all D-zoned, black-owned, homes) from a grade of D to a grade of C. Conversely, given that the share black within 100 meters on the right side of border was much smaller than it

was at distances beyond 100 meters into the D zone, moving all boundaries inward 100 meters would move an even larger number (and share) of white households out of the D zone. Specifically, this change in boundary location would shift 22,980 white households (37% of all D-zoned white households) and 14,007 white-owned homes (36% of all D-zoned, white-owned, homes) from a grade of D to a grade of C.

Of course, race-neutral zone assignment would not have implied that all boundaries would shift in by 100 meters. Instead, it would likely imply that only the specific boundaries that gave rise to the bump in percent black within the first 100 meters of the D zone would be shifted. Thus, the 20.9 percent reduction in black household redlining that would have resulted if all black individuals living along the D-side of the boundary were moved to C zones likely overestimates the impact of race-based boundary selection on the exposure of black households to redlining.

This observation motivates our second counterfactual, which we believe to be a more realistic exercise. Here, we compute the impact of the observed clustering under the assumption that, without racial bias in the choice of boundary location, we would have seen a smooth rise in probability beginning about 140 meters on the C-rated side of the C-D border and continuing about 360 meters on the D-rated side before accelerating. We thus consider the impact of relocating only the bump itself. We move the excess black households represented by the "bump" from an assignment of D to an assignment of C. This counterfactual exercise does not move any white households from the D zone to the C zone.

We start the process by fitting a nonlinear trend line to the data. We then then remove the positive deviation from trend that occurred within this band.[29] Under the exercise, 392 black

[29] See Appendix IV for details.

households (2.8 percent of all D-zoned black households) and 93 black-owned homes (2 percent of all D-zoned black-owned homes) are moved from a D security grade to a C security grade. Thus, this more nuanced counterfactual suggests a much smaller role for racially motivated distortions in the boundary location in explaining the overall exposure of black households to redlining.

The notion that race played a role in the choice of specific boundary locations is consistent with the importance placed on race by real estate professionals and scholars of the day, many of whom played a central role in the development of the HOLC maps. However, our analysis suggests that this focus on race was likely responsible for the redlining of only a modest number of black households. The black households that were redlined because of these boundary adjustments represent only a fraction of the overall number of the black families who found themselves redlined. The majority, more than 80 percent and quite possibly more than 95 percent, were redlined in the process of HOLC map making because they had few choices outside northern cities' most economically disadvantaged neighborhoods. Put differently, our results suggest that the majority of black households were redlined due to a combination of discrimination-driven economic disadvantage, direct discrimination in housing markets and discrimination in the provision of city services, all of which predated the creation of HOLC and the HOLC maps. These forces left black families and individuals with little choice but to live in neighborhoods that were destined to be redlined due to market conditions and their overall level of economic distress.

V. Conclusion

During the 1930s, the Home Owners Loan Corporation (HOLC) created a series of maps designed to summarize spatial variation in the riskiness of housing assets in different neighborhoods. These HOLC maps, in conjunction with contemporaneous maps produced by the Federal Housing Agency (FHA), are at the center of debates regarding the long-run impacts of government-imposed redlining. These maps are particularly salient because black households were almost entirely concentrated in the highest risk zones on these maps. This concentration, combined with the fact that these formerly redlined neighborhoods largely remain economically distressed today, has led many scholars to conclude that racial bias in the construction of the maps has had important effects over the long run.

Using newly digitized data for ten major northern cities, we assessed the HOLC maps for the importance of racially motivated boundary construction in explaining the prevalence of black residents in redlined neighborhoods. We found that the HOLC map boundaries were drawn in such a way as to capture pre-existing discontinuities in neighborhood economic characteristics and that racial bias in the construction of the HOLC maps can explain at most a small fraction of the observed concentration of black individuals in redlined zones. We also discussed the limitations of using the HOLC maps to study the long-term impacts of redlining activities undertaken by the FHA.

In sum, our results suggest that the vast majority of black households were redlined, not due to biased map construction, but instead because decades of disadvantage and discrimination had already pushed them in to the core of economically distressed neighborhoods prior to the government's direct involvement in mortgage markets. As a result, the HOLC maps are best viewed as providing clear evidence of how decades of unequal treatment effectively limited where black households could live in the 1930s rather than reflecting racial bias in the

construction of the maps themselves. As such, one contribution of our work is to further document the level of disadvantage experienced by black individuals in early twentieth century American cities – disadvantages that went unaddressed by government at any level until the 1960s.

Finally, our results suggest two important but difficult-to-answer questions. First, what could have been accomplished if, instead of adopting market norms, the FHA had leveraged its market involvement to actively push back against the patterns of economic and racial segregation that are documented in the HOLC maps? Second, setting the maps aside, what role did FHA policies play in locking in place economic challenges in the neighborhoods that were redlined by HOLC? We believe an initial answer to this second question comes directly from the former head of underwriting at the FHA, Frederick Babcock. Writing in 1939, Babcock (1939) focused on two key areas when summarizing his view on the FHA's impact. First, he noted that "valuation practice has become uniform," and second, that lenders now focus on "location before anything else" when considering real estate loans. To paraphrase Babcock, the FHA was responsible for creating a new unified approach to real estate lending which had as its central tenet an aversion to lending in the types of neighborhoods that were colored red on the HOLC maps. Assessing the importance of this policy innovation is a key question for future work on the role the federal government played in segregating American cities.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder. *The effects of the 1930s HOLC" redlining" maps*. No. 2017-12. Working Paper, 2017.

Akbar, Prottoy A., et al. *Racial segregation in housing markets and the erosion of black wealth*. No. w25805. National Bureau of Economic Research, 2019.

Aneja, Abhay P., and Carlos F. Avenancio-Leon. "The effect of political power on labor market inequality: Evidence from the 1965 Voting Rights Act." *Unpublished working paper* (2019).

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

Babcock, Frederick Morrison. *The valuation of real estate*. McGraw-Hill, 1932.

Brinkman, Jeffrey, and Jeffrey Lin. "Freeway Revolts!" *Federal Reserve Bank of Philadelphia Working Papers* 19-29 (2019).

Burgess, Ernest W. "Residential segregation in American cities." *The annals of the american academy of political and social science* 140.1 (1928): 105-115.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. "Robust data-driven inference in the regression-discontinuity design." *The Stata Journal* 14.4 (2014): 909-946.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82.6 (2014): 2295-2326.

Carruthers, Celeste K., and Marianne H. Wanamaker. "Separate and unequal in the labor market: human capital and the jim crow wage gap." *Journal of Labor Economics* 35.3 (2017): 655-696.

Coates, Ta-Nehisi. "The case for reparations." *The Atlantic* 313.5 (2014): 54-71.

Colin Gordon. *Mapping Decline: St. Louis and the Fate of the American City.* Philadelphia, University of Pennsylvania Press, 2008, 92.

Cressey, Paul Frederick. "Population succession in Chicago: 1898-1930." *American Journal of Sociology* 44.1 (1938): 59-69.

Federal Housing Administration. 1939. Fifth Annual Report of the Federal Housing Administration. Washington, D.C.: Government Printing Office.

Federal Housing Administration. 1941. Seventh Annual Report of the Federal Housing Administration. Washington, D.C.: Government Printing Office.

Federal Housing Administration. 1947. Thirteenth Annual Report. Washington: Government Printing Office.

Fishback, Price. 2014. "Saving the Neighborhood: Complementary Insights from Housing Markets in the 1920s and 1930s." *Arizona Law Review: The Syllabus* 56(3) (2014), online.

Fishback, Price, Sebastian Fleitas, Jonathan Rose, Kenneth Snowden, 2017, "Mortgage Foreclosure Overhangs and the Slow Recovery During the 1930s," Working Paper.

Fishback, Price, Alfonso Flores-Lagunes, William Horrace, Shawn Kantor, and Jaret Treber, 2011, "The Influence of the Home Owners' Loan Corporation on Housing Markets During the 1930s," *Review of Financial Studies* 24, 1782-1813.

Fishback, Price and Trevor Kollmann. "New Multi-City Estimates of the Changes in Home Values, 1920-1940." *Housing and Mortgage Markets in Historical Perspective*. Edited by White, Eugene, Kenneth Snowden, and Price Fishback. National Bureau of Economic Research Conference Volume. Chicago: University of Chicago Press, 2014, pp. 203-244.

Fishback, Price, Jonathan Rose, and Kenneth Snowden. *Well Worth Saving: How the New Deal Safeguarded Home Ownership*. Chicago: University of Chicago Press, 2013.

Fleitas, Sebastian, Price Fishback, and Kenneth Snowden. "Economic crisis and the demise of a popular contractual form: Building & Loans in the 1930s." *Journal of Financial Intermediation* 36 (2018): 28-44.

Freund, David MP. *Colored property: State policy and white racial politics in suburban America*. University of Chicago Press, 2010.

Gibbard, Harold A. "The Status Factor in Residential Successions." *American Journal of Sociology* 46.6 (1941): 835-842.

Greer, James. 2012. "The Home Owners' Loan Corporation and the Development of the Residential Security Maps." *Journal of Urban History* (39(2): 275-296.

Hillier, Amy E. "Redlining and the home owners' loan corporation." *Journal of Urban History* 29.4 (2003): 394-420.

Hillier, Amy E. "Who received loans? Home Owners' Loan Corporation lending and discrimination in Philadelphia in the 1930s." *Journal of Planning History* 2.1 (2003): 3-24.

Jackson, Kenneth T. *Crabgrass Frontier: The Suburbanization of the United States*. Oxford University Press, 1985.

Kollman, Trevor. 2011. Housing Markets, Government Programs, and Race During the Great Depression (2011). Unpublished Ph.D. dissertation, University of Arizona.

Krimmel, Jacob. "Persistence of prejudice: Estimating the long term effects of redlining." (2018).

Kucheva, Yana, and Richard Sander. "The misunderstood consequences of Shelley v. Kraemer." *Social science research* 48 (2014): 212-233.

LaVoice, Jessica. *Essays about Race, Discrimination, and Inequality* (2020). Unpublished Ph.D dissertation, University of Pittsburgh.

Long, Jason, and Joseph Ferrie. "Intergenerational occupational mobility in Great Britain and the United States since 1850." *American Economic Review* 103.4 (2013): 1109-37.

Massey, Douglas, and Nancy A. Denton. *American apartheid: Segregation and the making of the underclass*. Harvard university press, 1993.

Michney, Todd and LaDale Winling. 2019. "New Perspectives on New Deal Housing Policy: Explicating and Mapping HOLC Loans to African Americans." *Journal of Urban History*: 1-31.

National Archives. Undated. "A Review of the Objectives, Achievements, Administration Operations, and Related Policies." In Finding Aid for the Home Owners' Loan Corporation City Survey Program." Record Group 39. Federal Home Loan Bank Board. National Archives II in College Park Maryland. Xerox. Unpublished.

Rose, Jonathan D. "The incredible HOLC? Mortgage relief during the Great Depression." *Journal of Money, Credit and Banking* 43.6 (2011): 1073-1107.

Rose, Jonathan D., and Kenneth A. Snowden. "The New Deal and the origins of the modern American real estate loan contract." *Explorations in Economic History* 50.4 (2013): 548-566.

Rothstein, Richard. 2017. *The Color of Law: A Forgotten History of How our Government Segregated America.* New York: Liveright Publishing Corporation.

Saavedra, Martin, and Tate Twinam. "A machine learning approach to improving occupational income scores." *Explorations in Economic History* 75 (2020): 101304.

Satter, Beryl. *Family properties: Race, real estate, and the exploitation of black urban America*. Macmillan, 2009.

Schietinger, E. F. "Racial succession and value of small residential properties." *American Sociological Review* 16.6 (1951): 832-835.

Schietinger, Egbert F. "Race and residential market values in Chicago." *Land Economics* 30.4 (1954): 301-308.

Shertzer, Allison, Randall P. Walsh, and John R. Logan. "Segregation and neighborhood change in northern cities: New historical GIS data from 1900–1930." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 49.4 (2016): 187-197.

Snowden, Kenneth A. "Building and loan associations in the US, 1880–1893: The origins of localization in the residential mortgage market." *Research in Economics* 51.3 (1997): 227-250.

Snowden, Kenneth. "Mortgage banking in the United States, 1870-1940." Research Institute for Housing America Research Paper 13-02 (2013).

Snowden, Kenneth A. "The transition from Building and Loan to Savings and Loan, 1890-1940." *Finance, intermediaries and economic development* (2003): 157-208.

Sugrue, Thomas J. *The origins of the urban crisis: Race and inequality in postwar detroit-updated edition*. Princeton University Press, 2014.

Wilder, Craig. *A Covenant with Color: Race and Social Power in Brooklyn. New York*. Columbia University Press, 2000.

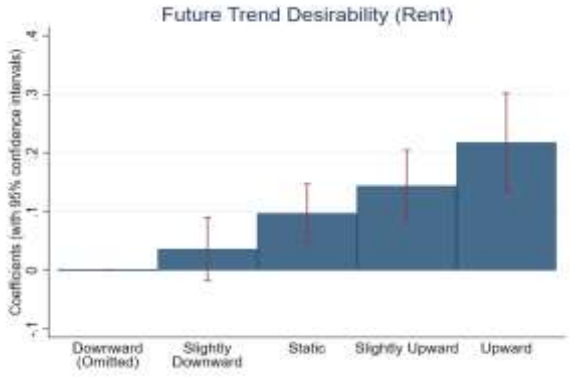
U.S. Bureau of Census. *Sixteenth Census of the U.S- 1940 Housing: Volume III-Mortgages on Owner-Occupied Non Farm Homes: Part 1: United States Summary*. Washington, DC: Government Printing Office, 1943.

Van Dyke, W. S. "A NEW PLAN FOR HOME PURCHASE." Bankers' Magazine (1896-1943) 118.6 (1929): 934.

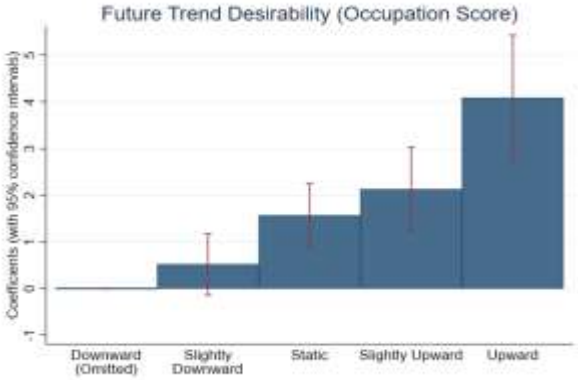
Figure 1: Coefficients and Confidence Intervals from Zone-Level Regressions of the Change between 1930 and 1940 in Census Outcomes as a Function of HOLC Reports of Future Trend Desirability



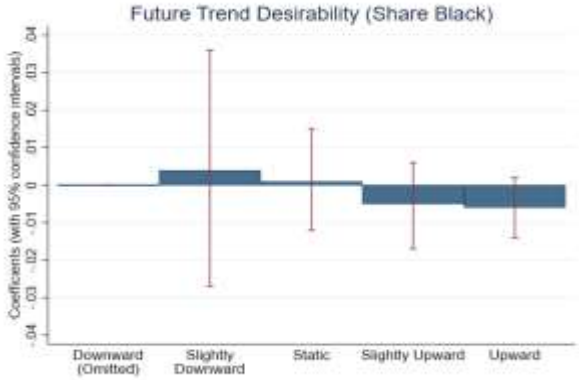
Panel A: House Values



Panel B: Rents



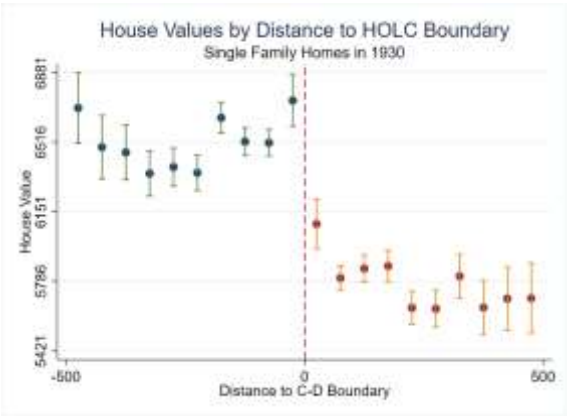
Panel C: Occupation Scores



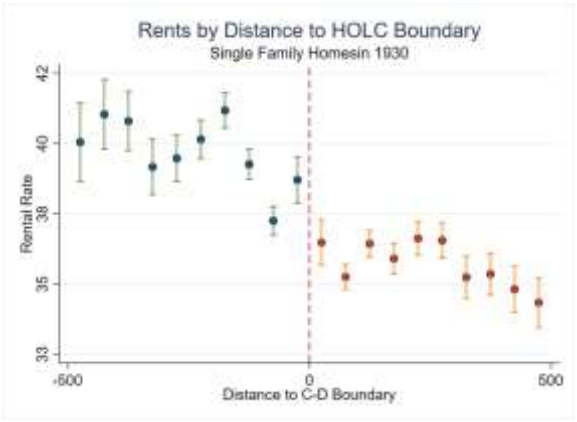
Panel D: Racial Composition

Notes: This figure shows the coefficient and 95% confidence interval for the indicator variables related to the future trend desirability of a neighborhood with downward being the omitted category. Each regression controls for the 1930 value of the outcome variable. Regression results are also reported in Appendix Table A3

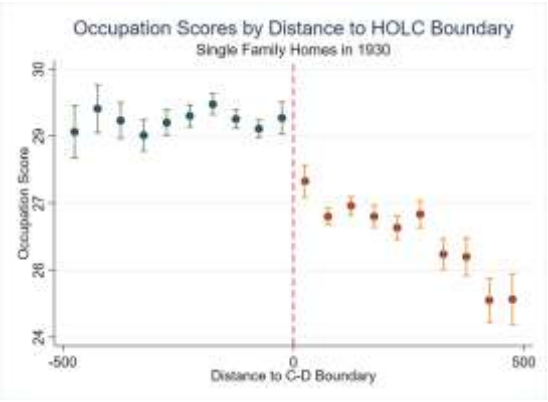
Figure 2: 1930 Levels by Distance in 50-Meter Bins to HOLC Boundary from 1930 Census Address Data



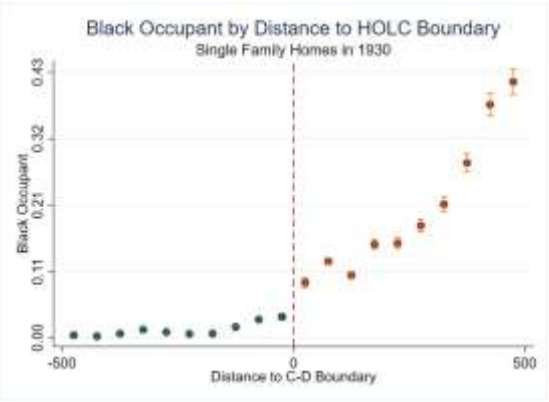
Panel A: House Values



Panel B: Rents



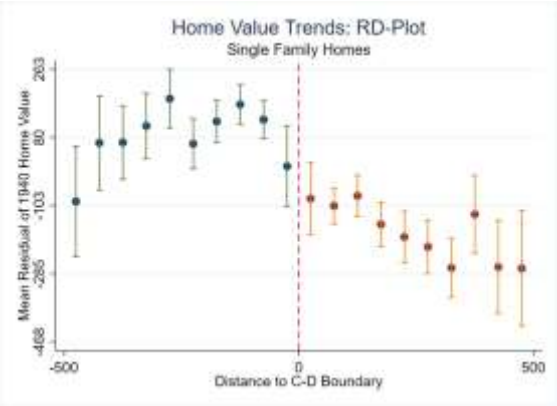
Panel C: Occupation Scores



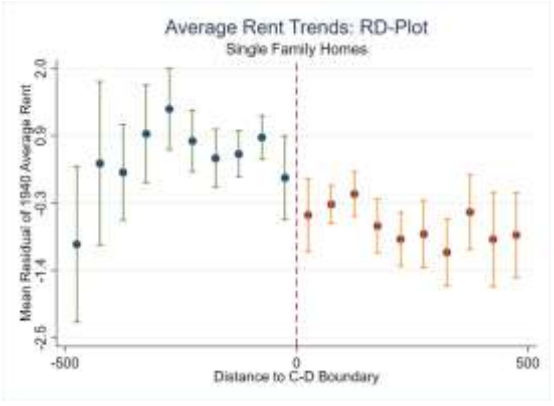
Panel D: Racial Composition

Notes: This figure shows averages of 1930 census data for single family households from 50-meter bins by distance to a C-D HOLC boundary. The red dotted line represents the HOLC boundary, positive distances represent households in the redlined zone, and negative distances represent houses in the yellow-lined zones. All distances are measured in meters.

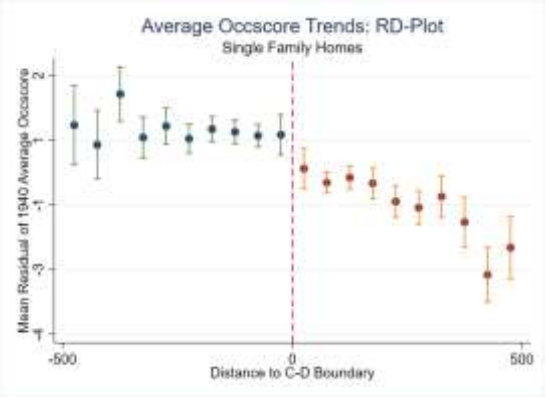
Figure 3: Changes between 1930 and 1940 by Distance to HOLC Boundary from Census Address Data



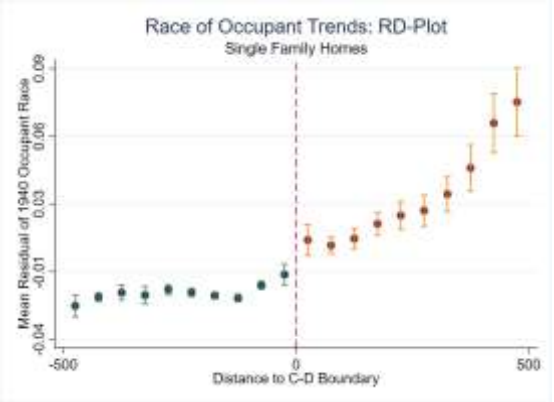
Panel A: House Values



Panel B: Rents



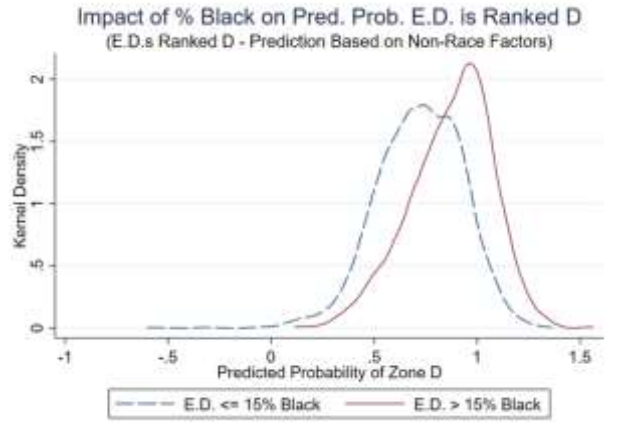
Panel C: Occupation Scores



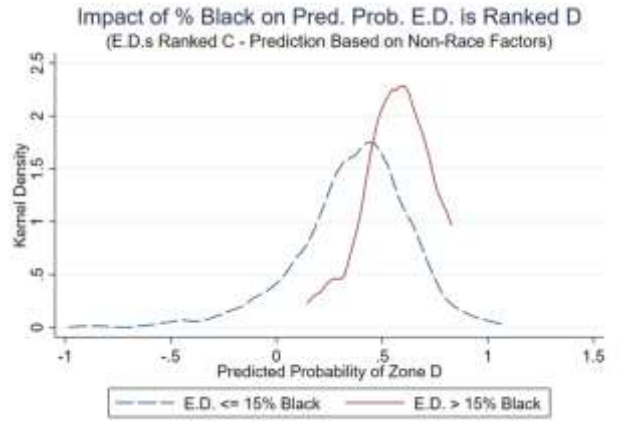
Panel D: Racial Composition

Notes: This figure shows averages of 1940 census data, controlling for 1930 values, for single family households from 50-meter bins by distance to a C-D HOLC boundary. The red dotted line represents the HOLC boundary, positive distances represent households in the redlined zone, and negative distances represent houses in the yellow-lined zones. All distances are measured in meters.

Figure 4: Distributions of Predicted Probability the Enumeration District (ED) is Rated D Based on Other Characteristics When Percent Black is Above or Below 15 Percent



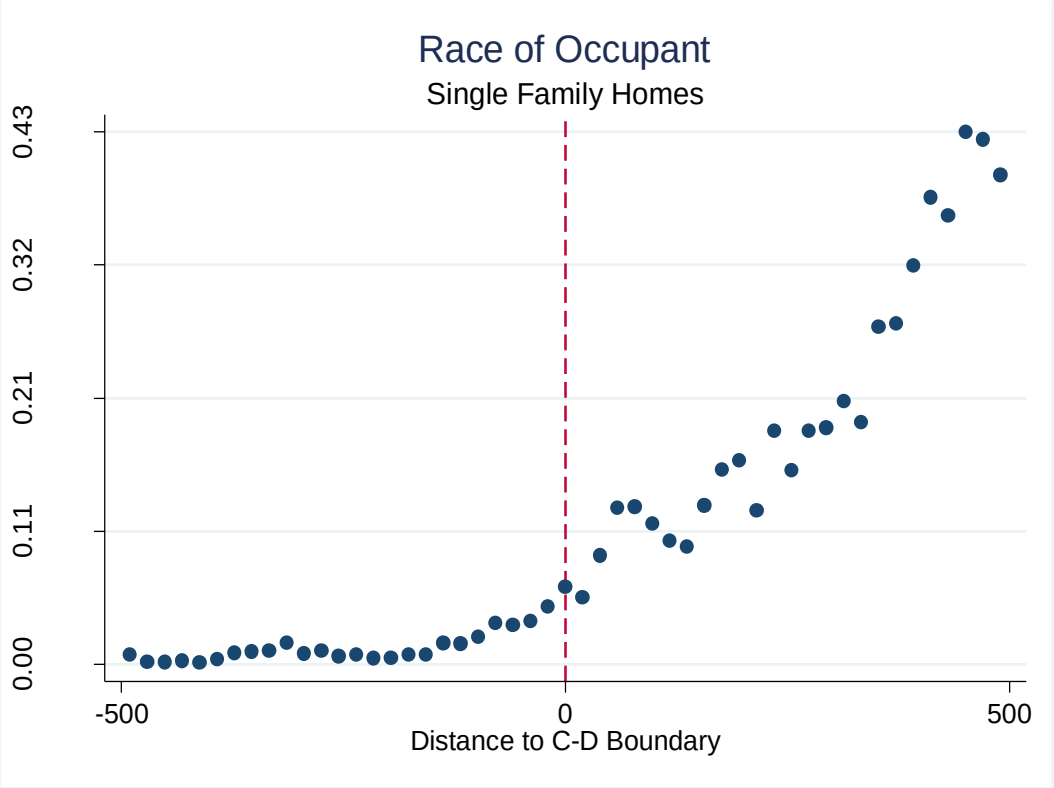
Panel A: D Zones



Panel B: C Zones

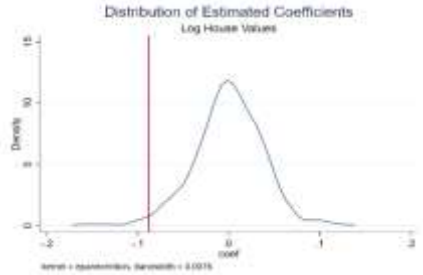
Notes: This figure shows the distribution of an enumeration districts predicted probability of being redlined. Predicted probabilities were calculated using 1930 ED census data from regression equation 3 in which a dummy with value 1 for Zone D is regressed on the share foreign born, share homeowners, average age, average occupation score, average rent, average sales price, and labor force participation rate. Data are for EDs with the share black greater than 15 percent and less than or equal to 15 percent.

Figure 5: Share of Black Families in 20-Meter Bins by Distance to C-D Boundaries

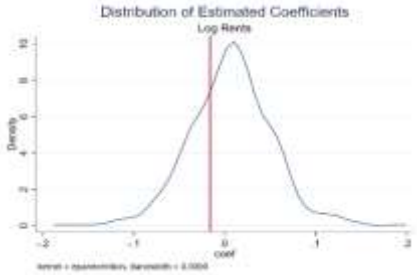


Notes: This figure shows the relationship between race and distance to a C-D boundary. It shows the share of black residents based on distance to a HOLC boundary, with negative distances representing locations on the C side of the boundary and positive distances representing areas on the D side of the boundary.

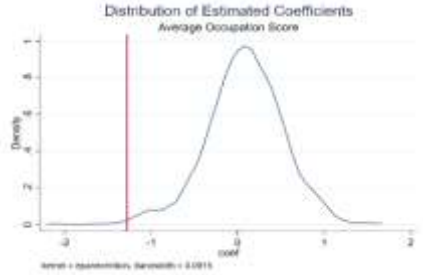
Figure 6: R-D Coefficient Estimates from Actual HOLC Boundary in Table 2 (Red Line) Compared with Distribution of Coefficients When Locations of Boundary Are Randomly Assigned 1,000 Times



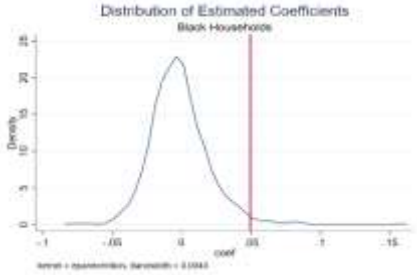
Panel A: House Values 1930 (p=0.016)



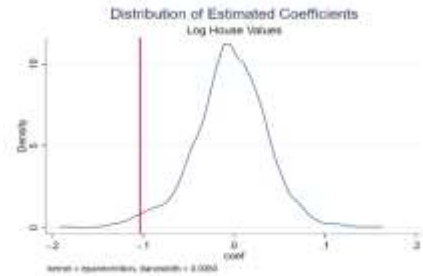
Panel B: Rents 1930 (p=0.281)



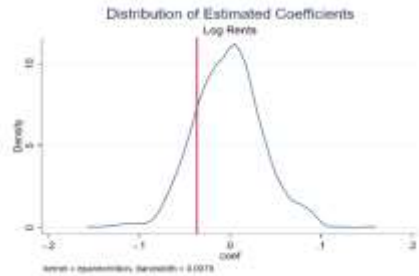
Panel C: Occupation Scores 1930 (p=0.005)



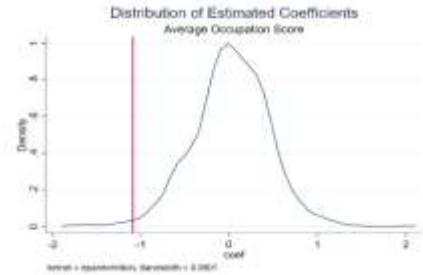
Panel D: Share Black 1930 (p=0.018)



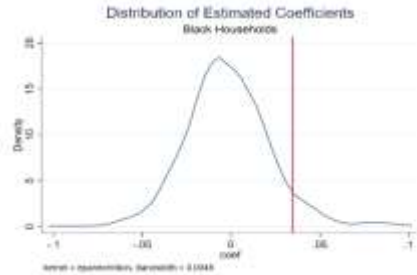
Panel E: House Values 1940 (p==0.016)



Panel F: Rents 1940 (p=0.166)



Panel G: Occupation Scores 1940 (p=0.011)



Panel H: Share Black 1940 (p=0.060)

Notes: The vertical red line shows the value of the coefficient associated with the actual HOLC boundary from the R-D analysis in Table 2. The blue line shows the empirical distribution of counterfactual coefficient estimates from randomly chosen boundary locations (1000 replications). Counterfactual boundaries vary in 5-meter increments from 250 meters on the C-side of actual boundaries to 250 meters into the D-side. One-tailed empirical p-values for the estimate at true borders are also reported.

Table 1: 1930 Census and HOLC Survey Data Summary Statistics by Zone

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Zone A All	Zone B All	Zone C All	Zone D All	Zone C Hi-Black	Zone C Lo-Black	Zone D Hi-Black	Zone D Low-Black
Panel A: HOLC Survey Data (Zone)								
Family Income	21716	6836	3241	1467	3315	3236	1343	1563
	(22467)	(8453)	(6242)	(827)	(2404)	(6457)	(599)	(969)
House Value	23356	9042	5776	3799	5681	5779	3435	4005
	(18059)	(6022)	(2637)	(2269)	(2677)	(2641)	(2308)	(2230)
Rent	67.55	50.00	36.38	20.46	36.42	36.38	18.88	21.40
	(26.10)	(12.98)	(12.05)	(8.70)	(17.28)	(11.85)	(8.59)	(8.67)
Share Black	0.00	0.00	0.00	0.19	0.05	0.00	0.48	0.02
	(0.00)	(0.00)	(0.02)	(0.30)	(0.05)	(0.02)	(0.33)	(0.05)
Share Foreign	0.02	0.09	0.27	0.45	0.13	0.27	0.32	0.51
	(0.06)	(0.16)	(0.27)	(0.27)	(0.17)	(0.27)	(0.23)	(0.27)
Share of Houses Occupied	0.743	0.8779	0.9452	0.7889	0.853	0.9483	0.7501	0.8096
	(0.44)	(0.30)	(0.17)	(0.36)	(0.32)	(0.16)	(0.39)	(0.34)
Share Owner Occupied	0.66	0.701	0.6469	0.398	0.375	0.6561	0.2808	0.4639
	(0.41)	(0.29)	(0.23)	(0.30)	(0.25)	(0.22)	(0.22)	(0.32)
Future Trend Desirability								
Upward	0.368	0.134	0.027	0	0.1	0.022	0	0
	(0.50)	(0.34)	(0.16)	(0)	(0.316)	(0.15)	(0)	(0)
Slightly Upward	0.158	0.254	0.027	0.034	0	0.029	0	0.056
	(0.38)	(0.44)	(0.16)	(0.18)	(0.00)	(0.17)	(0)	(0.23)
Static	0.474	0.373	0.231	0.333	0.5	0.212	0.283	0.366
	(0.51)	(0.49)	(0.42)	(0.47)	(0.53)	(0.41)	(0.46)	(0.49)
Slightly Downward	0	0.104	0.095	0.077	0.2	0.088	0.065	0.085
	(0)	(0.31)	(0.30)	(0.27)	(0.42)	(0.28)	(0.25)	(0.28)
Downward	0	0.134	0.619	0.556	0.2	0.65	0.652	0.493
	(0)	(0.34)	(0.49)	(0.50)	(0.42)	(0.48)	(0.48)	(0.50)

Table 1, con't

Panel B: 1930 Census (Zone)								
Occupation Score	36.87	31.00	27.41	23.73	27.209	27.419	22.607	24.35
	(5.37)	(4.47)	(3.09)	(1.90)	(4.93)	(3.02)	(1.63)	(1.76)
House Value	10411.40	9733.96	8358.17	7051.11	8525.53	8352.035	6720.08	7236.995
	(3156.85)	(3189.17)	(2611.35)	(2977.47)	(3055.96)	(2599.96)	(2966.81)	(2978.69)
Rent	58.55	55.93	46.865	32.33	43.954	46.97	32.486	32.24
	(16.85)	(15.42)	(13.76)	(10.47)	(6.01)	(13.96)	(8.78)	(11.33)
Share Black	0.017	0.004	0.006	0.121	0.092	0.003	0.326	0.006
	(0.022)	(0.007)	(0.019)	(0.223)	(0.041)	(0.005)	(0.271)	(0.009)
Share Foreign	0.204	0.179	0.239	0.283	0.193	0.241	0.225	0.316
	(0.123)	(0.088)	(0.088)	(0.121)	(0.10)	(0.09)	(0.14)	(0.10)
Share Labor Force	0.494	0.441	0.453	0.453	0.499	0.452	0.476	0.441
	(0.111)	(0.091)	(0.075)	(0.081)	(0.09)	(0.07)	(0.08)	(0.08)
Share Owner Occupied	0.437	0.442	0.405	0.288	0.324	0.408	0.263	0.302
	(0.288)	(0.238)	(0.196)	(0.190)	(0.181)	(0.197)	(0.185)	(0.192)
Total Homes Owned	201.15	931.51	1127.66	1450.17	809.00	1139.21	1949.85	1171.72
	(168.72)	(1491.58)	(1351.90)	(2213.44)	(657.17)	(1369.72)	(2946.67)	(1619.63)
Total Homes Owned Black	0.27	1.01	4.02	107.78	53.20	2.24	289.30	6.63
	(1.00)	(3.34)	(16.10)	(347.36)	(48.06)	(10.10)	(536.20)	(19.080)
Total Homes Owned White	200.85	930.21	1123.31	1341.32	755.4	1136.64	1658.88	1164.36
	(168.57)	(1489.75)	(1347.48)	(2081.94)	(626.86)	(1365.20)	(2720.66)	(1606.45)
Number of Zones	26	112	286	204	10	276	73	131
Number of EDs	68	725	2344	3791	92	2240	1952	1839
Number of People	79,551	1,046,617	3,484,700	5,410,078	122,639	3,362,061	2,759,526	2,650,552
White Population	78,272	1,042,099	3,462,193	4,459,905	112,564	3,349,629	1,835,016	2,624,889
Black Population	1,279	4,518	22,507	950,173	10,075	12,432	924,510	25,663
Number of Owned Homes	5,230	104,329	322,512	295,834	8,090	314,422	142,339	153,495
Houses Owned by Whites	5,222	104,184	321,267	273,629	7,554	313,713	121,098	152,531
Houses Owned by Blacks	7	113	1,151	21,987	532	619	21,119	868

Notes: Standard deviations are reported in parentheses below the means. Data in this table come from both aggregated census enumeration districts as well as HOLC survey data. We limit our sample to only enumeration districts that are at least 90% contained in a given zone. Black zones represent zones with an above average share of black residents and white zones represent zones with fewer than the average share of black residents of 4.3 percent.

Table 2: Levels and Trends at C-D Boundary from Regression Discontinuity Regressions Using 1930 and 1940 Census Addresses

	(1) 1930	(2) 1930	(3) 1940	(4) 1940	(5) Trends	(6) Trends
Panel A: Log House Values						
Redlined Side	-0.098***	-0.087***	-0.153***	-0.104***	-0.061***	-0.055***
	(0.014)	(0.013)	(0.017)	(0.016)	(0.021)	(0.012)
Panel B: Log Rents						
Redlined Side	-0.02	-0.016	-0.043***	-0.037**	-0.049	-0.02
	(0.019)	(0.017)	(0.018)	(0.017)	(0.029)	(0.029)
Panel C: Occupation Score						
Redlined Side	-1.435***	-1.279***	-1.123***	-1.091***	-0.856***	-0.798**
	(0.179)	(0.180)	(0.188)	(0.190)	(0.291)	(0.302)
Panel D: Share Black						
Redlined Side	0.064***	0.049***	0.039***	0.035***	0.013**	0.016***
	(0.007)	(0.006)	(0.01)	(0.009)	(0.005)	(0.005)
Optimal Bandwidth	Yes	Yes	Yes	Yes	Yes	Yes
Matched Sample	Yes	Yes	Yes	Yes	Yes	Yes
Boundary FE	No	Yes	No	Yes	No	Yes

Notes: Standard errors are in parentheses below the coefficient. Each coefficient is estimated from a separate regression. *p < .10, **p < .05, ***p < .01. We used the optimal bandwidth selection procedure proposed by Calonico, Cattaneo, and Titiunik (2014).

Appendix

I. Figures and Tables

Figure A1: Pittsburgh Homeowners Loan Corporation Map

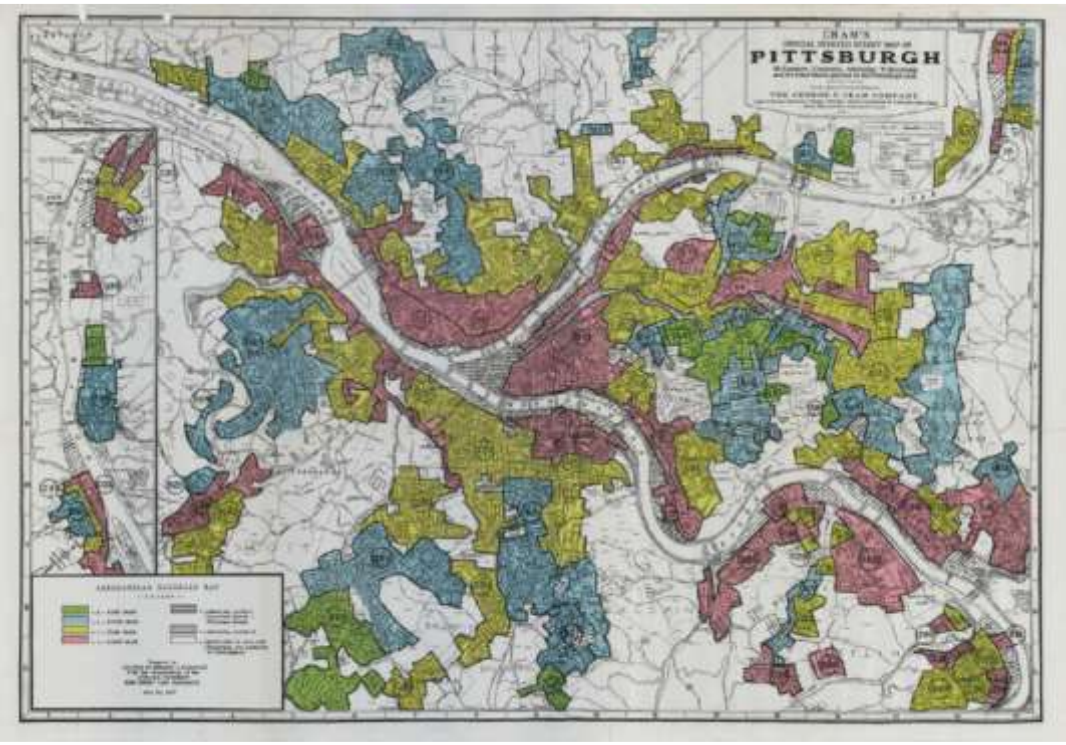
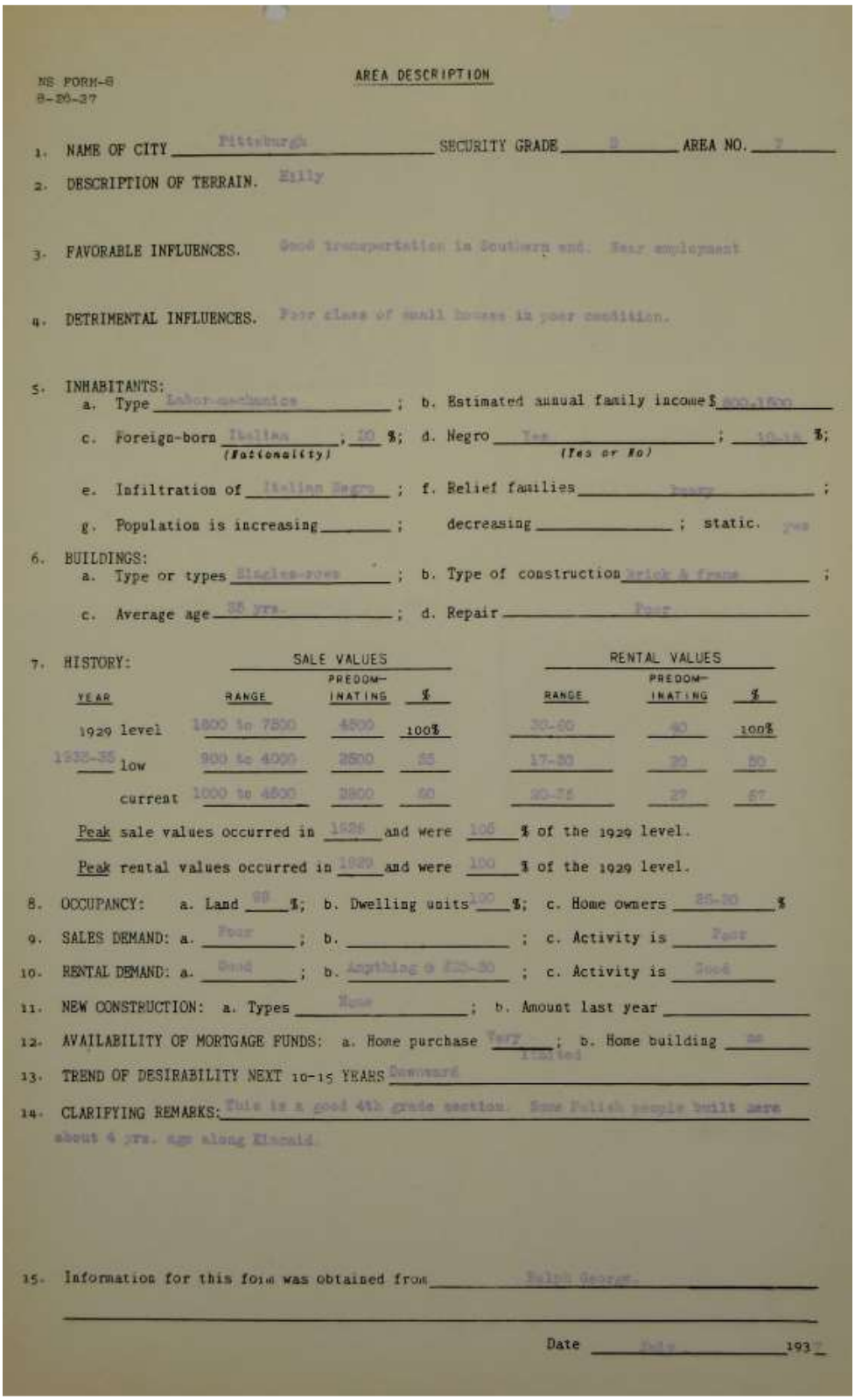


Figure A2: Example of HOLC Zone Survey (Pittsburgh D7 Zone)



NS FORM-8
8-20-37

AREA DESCRIPTION

1. NAME OF CITY Pittsburgh SECURITY GRADE D AREA NO. D7

2. DESCRIPTION OF TERRAIN. Hilly

3. FAVORABLE INFLUENCES. Good transportation in Southern end. Near employment

4. DETRIMENTAL INFLUENCES. Poor class of small houses in poor condition.

5. INHABITANTS:
a. Type Labor-mechanics ; b. Estimated annual family income $ 900-1600
c. Foreign-born Italian (Nationality), 20 %; d. Negro Yes (Yes or No); 10-15 %;
e. Infiltration of Italian Negro ; f. Relief families Heavy ;
g. Population is increasing ; decreasing ; static. yes

6. BUILDINGS:
a. Type or types Single-zone ; b. Type of construction Brick & frame ;
c. Average age 35 yrs. ; d. Repair Poor

7. HISTORY:

	SALE VALUES			RENTAL VALUES		
YEAR	RANGE	PREDOMINATING	%	RANGE	PREDOMINATING	%
1929 level	1800 to 7500	4500	100%	30-60	40	100%
1933-35 low	900 to 4000	2500	55	17-30	20	50
current	1000 to 4500	2800	60	20-75	27	67

Peak sale values occurred in 1926 and were 105 % of the 1929 level.

Peak rental values occurred in 1929 and were 100 % of the 1929 level.

8. OCCUPANCY: a. Land 98 %; b. Dwelling units 100 %; c. Home owners 85-90 %

9. SALES DEMAND: a. Poor ; b. ; c. Activity is Fair

10. RENTAL DEMAND: a. Good ; b. Anything @ $25-30 ; c. Activity is Good

11. NEW CONSTRUCTION: a. Types None ; b. Amount last year

12. AVAILABILITY OF MORTGAGE FUNDS: a. Home purchase Very limited ; b. Home building

13. TREND OF DESIRABILITY NEXT 10-15 YEARS Downward

14. CLARIFYING REMARKS: This is a good 4th grade section. Some Polish people built here about 4 yrs. ago along Elmerald.

15. Information for this form was obtained from Ralph George.

Date July 193?

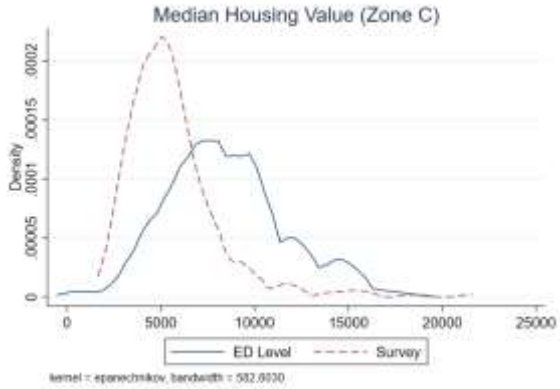
Figure A3: Distributions of HOLC Survey Mid-point and 1930 Median Census Housing Values



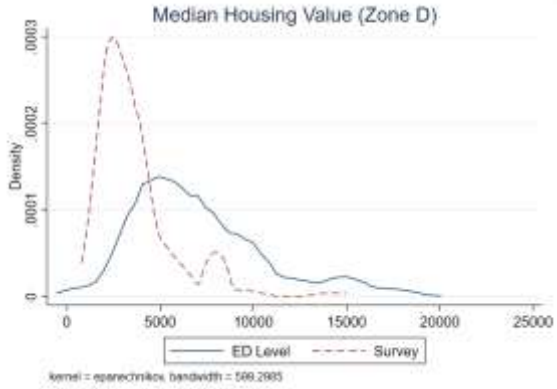
Panel A: Zone A



Panel B: Zone B



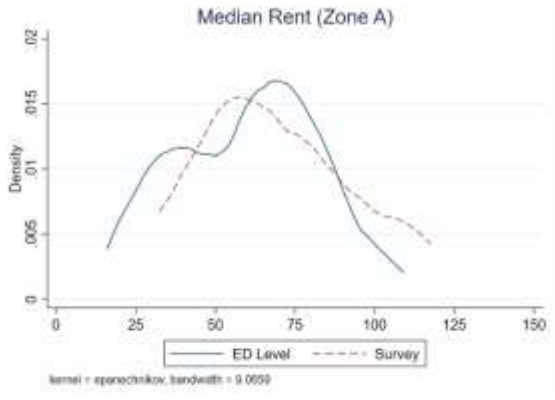
Panel C: Zone C



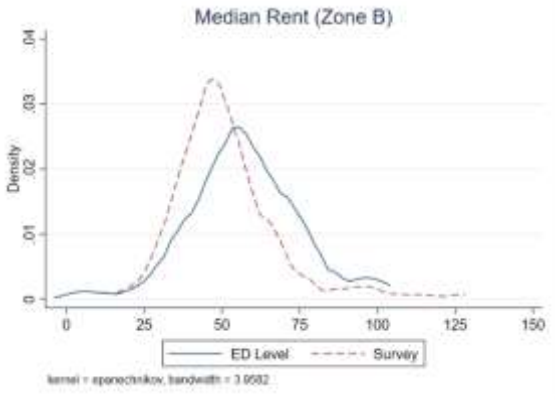
Panel D: Zone D

Notes: To analyze the accuracy of the surveys, we compare the kernel densities of survey measures of housing values with census measures of housing values. We use enumeration districts as our unit of observation for the census data and restrict our sample to only enumeration districts that lie completely within a residential security zone to prevent bias in our estimates.

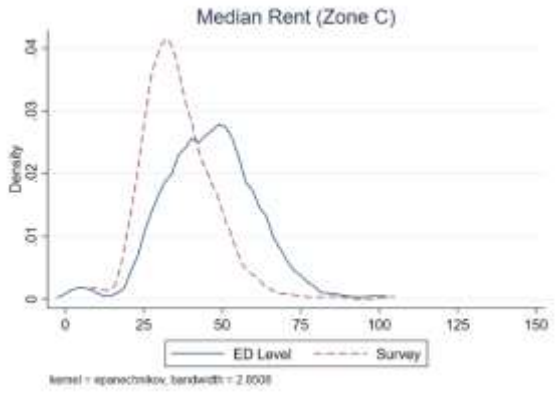
Figure A4: Distributions of HOLC Survey Mid-point and 1930 Census Median Rents



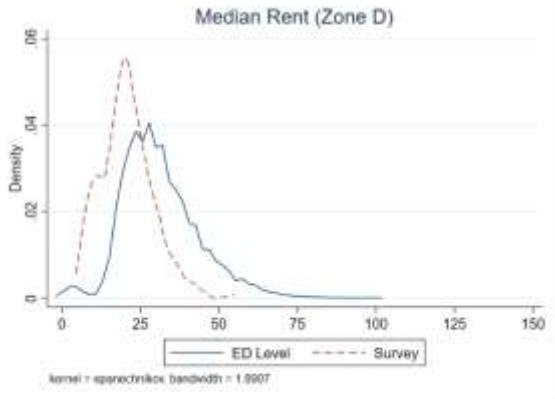
Panel A: Zone A



Panel B: Zone B



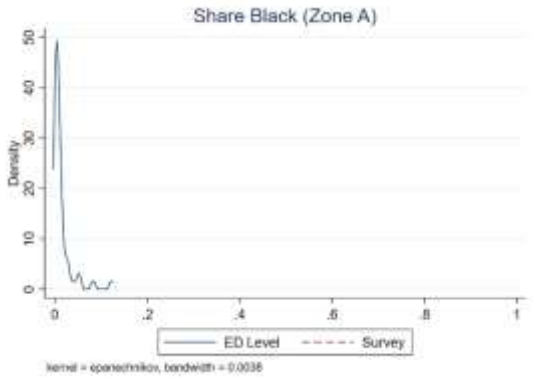
Panel C: Zone C



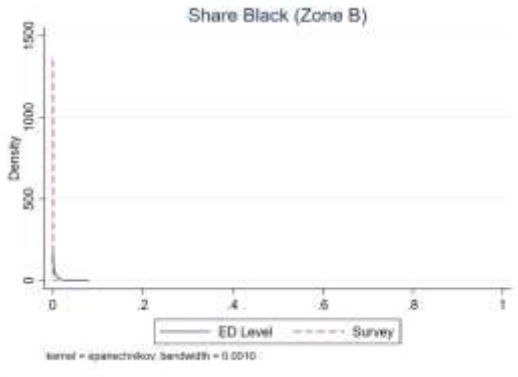
Panel D: Zone D

Notes: To analyze the accuracy of the surveys, we compare the kernel densities of survey measures and census measures of average rents. We use enumeration districts as our unit of observation for the census data and restrict our sample to only enumeration districts that lie completely within a residential security zone to prevent bias in our estimates.

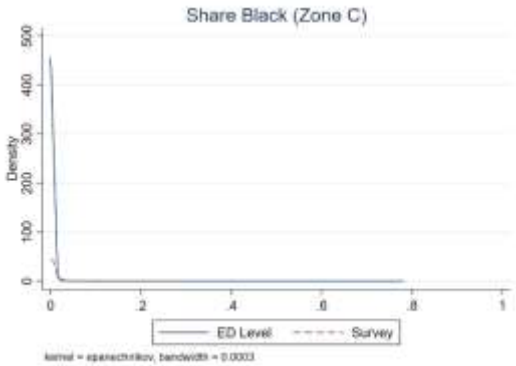
Figure A5: Distributions of HOLC Survey and 1930 Census Share Black



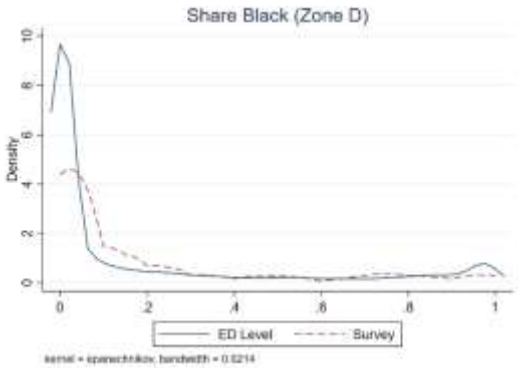
Panel A: Zone A



Panel B: Zone B



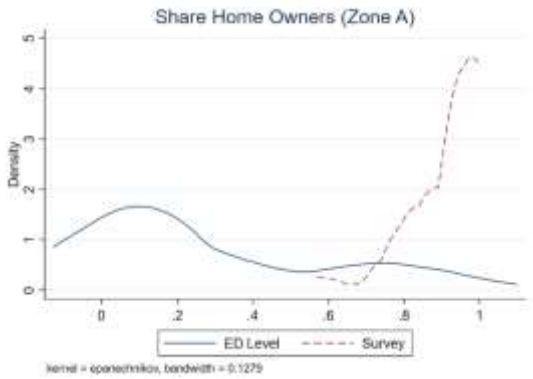
Panel C: Zone C



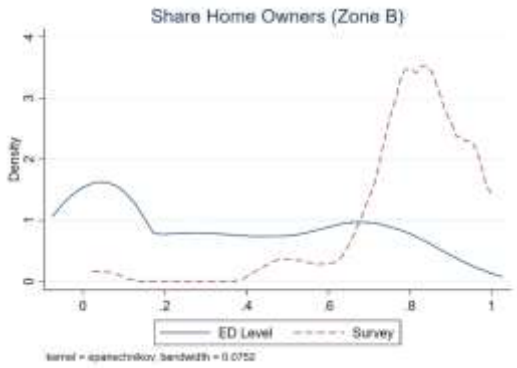
Panel D: Zone D

Notes: To analyze the accuracy of the surveys, we compare the kernel densities of survey and census measures of neighborhood racial composition. We use enumeration districts as our unit of observation for the census data and restrict our sample to only enumeration districts that lie completely within a residential security zone to prevent bias in our estimates.

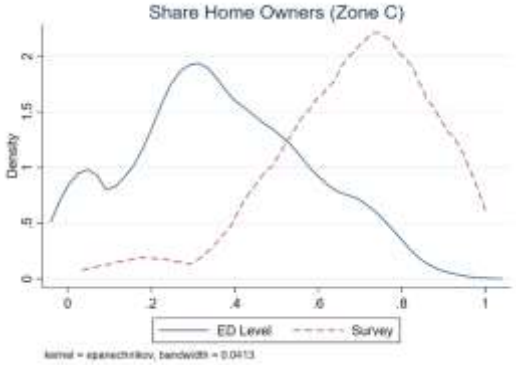
Figure A6: Distributions of HOLC Survey versus 1930 Census Home Ownership Rates



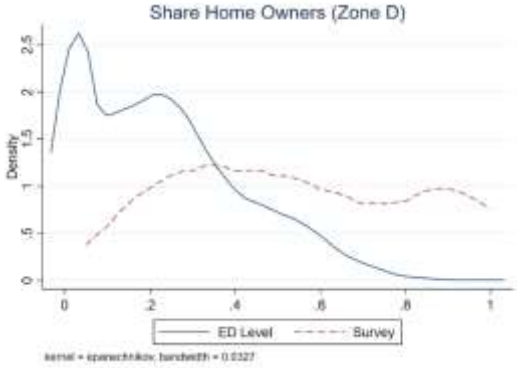
Panel A: Zone A



Panel B: Zone B



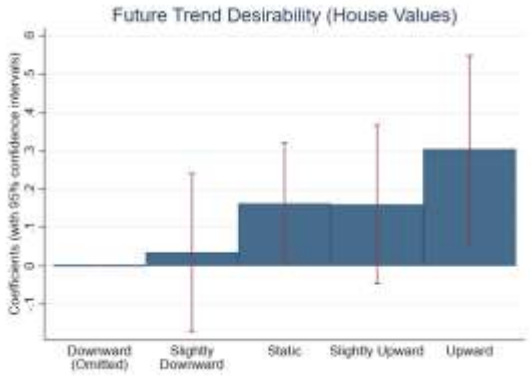
Panel C: Zone C



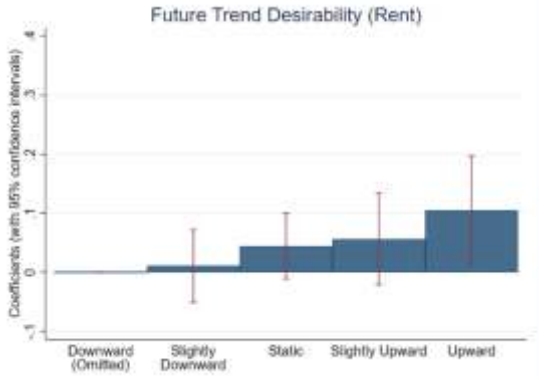
Panel D: Zone D

Notes: To analyze the accuracy of the surveys, we compare the kernel densities of survey and census measures of home ownership rates. We use enumeration districts as our unit of observation for the census data and restrict our sample to only enumeration districts that lie completely within a residential security zone to prevent bias in our estimates.

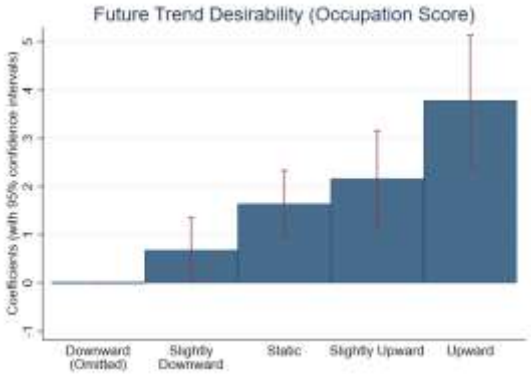
Figure A7: Coefficients and Confidence Intervals from Zone-Level Fixed-Effects Regressions of the Change between 1930 and 1940 in Census Outcomes as a Function of HOLC Reports of Future Trend Desirability



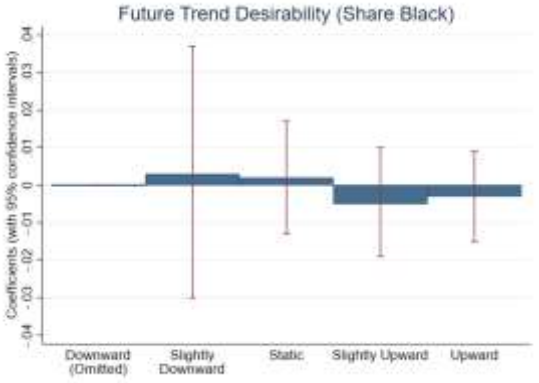
Panel A: House Values



Panel B: Rents



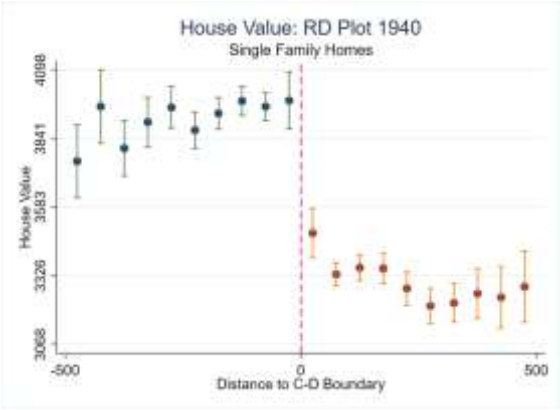
Panel C: Occupation Scores



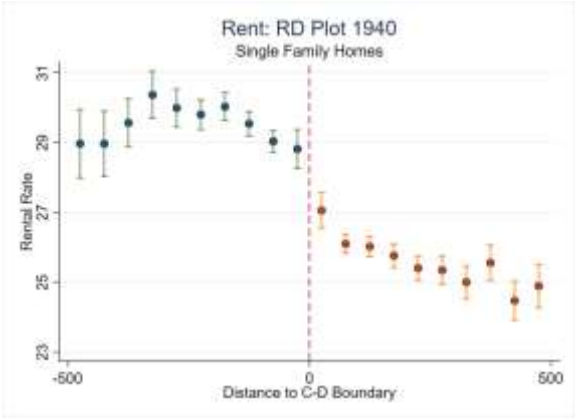
Panel D: Racial Composition

Notes: This figure shows the coefficient and 95% confidence interval for the indicator variables related to the future trend desirability of a neighborhood with downward being the omitted category. Each panel presents the results from a separate regression. Each regression controls for the 1930 value of the outcome variable and includes security grade fixed effects. Regression results are also reported in Appendix Table A3.

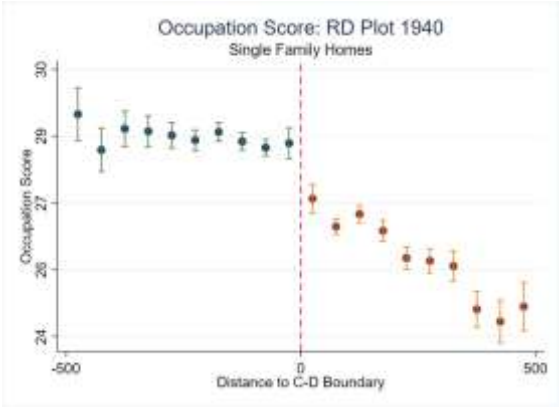
Figure A8: 1940 Census Levels by Distance in 50-Meter Bins to HOLC Boundary



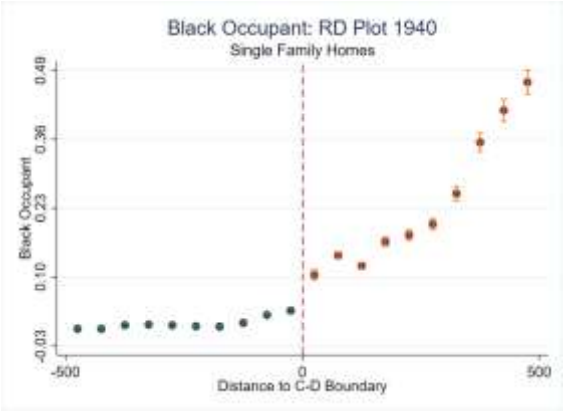
Panel A: House Values



Panel B: Rents



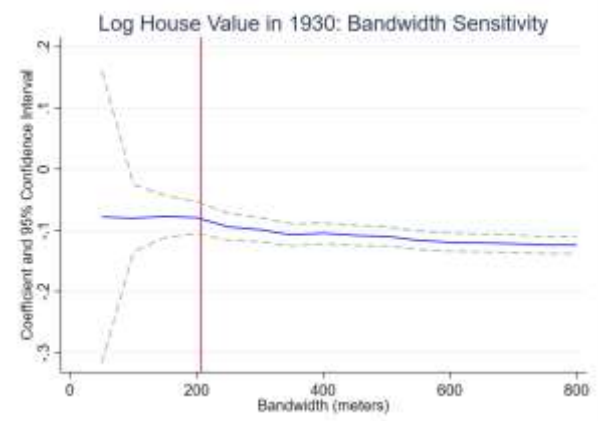
Panel C: Occupation Scores



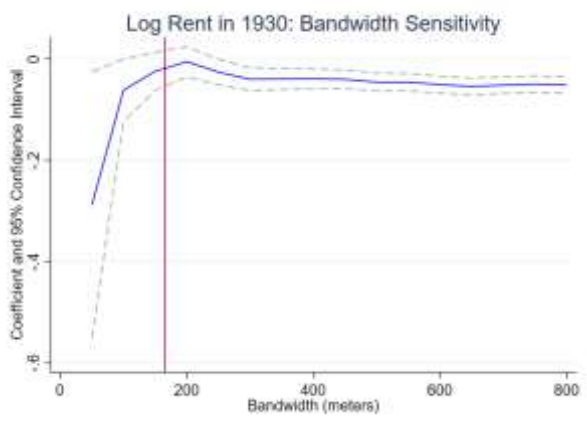
Panel D: Racial Composition

Notes: This figure shows binned averages of 1940 census data for single family households by distance in 50-meter bins to a CD HOLC boundary. Data within 30 meters (equivalent to the average depth of a household lot) are dropped to mitigate any measurement error. The red dotted line represents the HOLC boundary, positive distances represent households in the redlined zone, and negative distances represent houses in the yellow-lined zones. All distances are measured in meters.

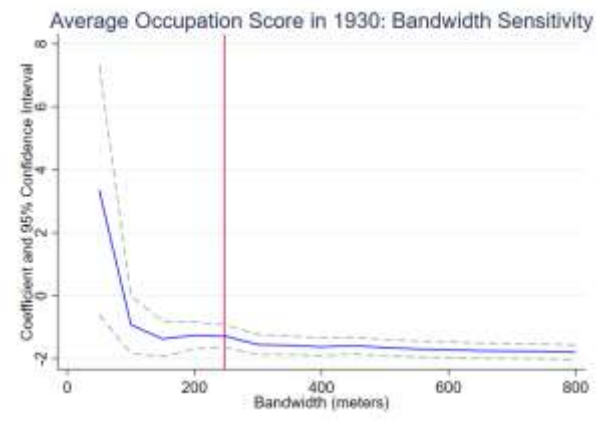
Figure A9: Optimal Bandwidth and Sensitivity of Results in Figure 2 to Different Bandwidths for the Boundary for 1930 Census Levels



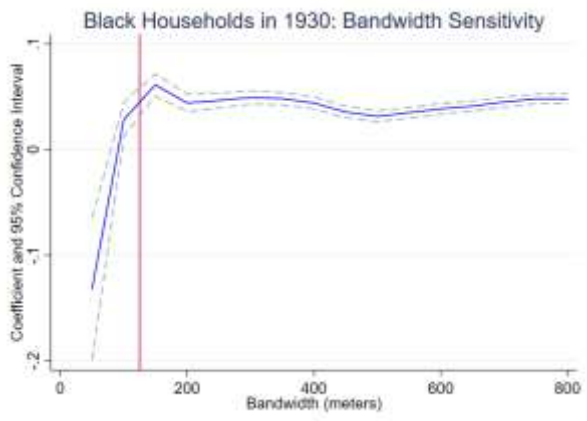
Panel A: House Values



Panel B: Rents



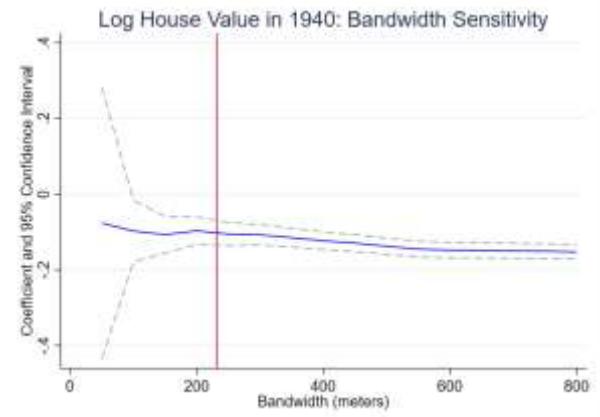
Panel C: Occupation Scores



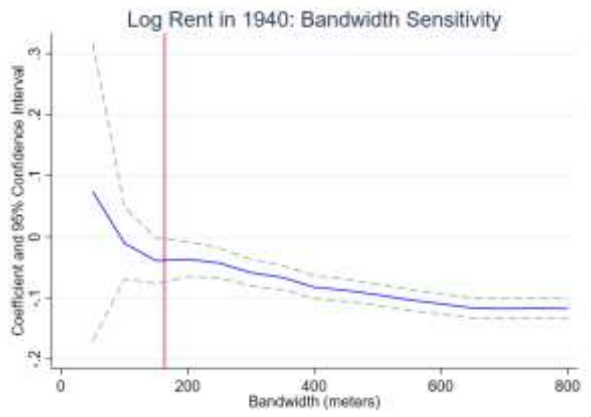
Panel D: Racial Composition

Notes: This figure graphs the estimated coefficients and 95% confidence intervals of β from equation (1) as we vary the bandwidth around a HOLC boundary. The red line represents the optimal bandwidth selected by the procedure proposed by Calonico, Cattaneo, and Titiunik (2014).

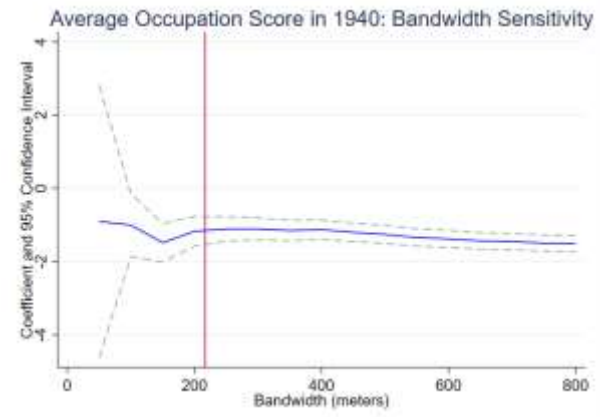
Figure A10: Optimal Bandwidth and Sensitivity of Results in Appendix Figure A8 to Different Bandwidths for the Boundary for 1940 Census Levels



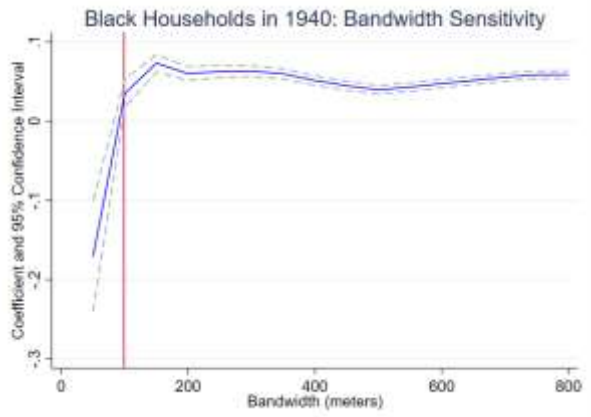
Panel A: House Values



Panel B: Rents



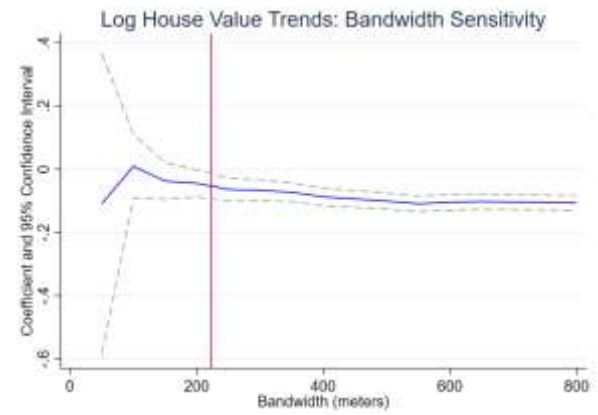
Panel C: Occupation Scores



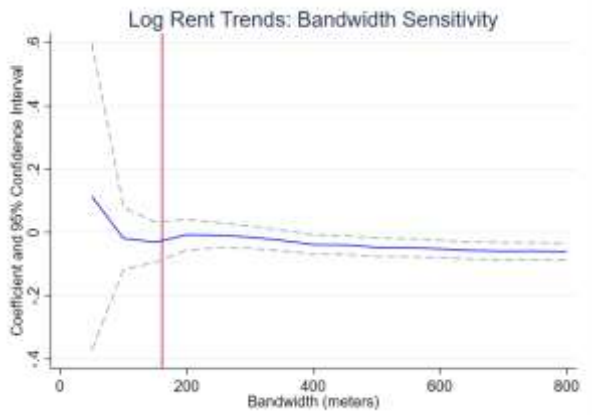
Panel D: Racial Composition

Notes: This figure graphs the estimated coefficients and 95% confidence intervals of β from equation (1) as we vary the bandwidth around a HOLC boundary. Bandwidth selection ranges from 50 to 800 meters. The red line represents the optimal bandwidth selected by the procedure proposed by Calonico, Cattaneo, and Titiunik (2014).

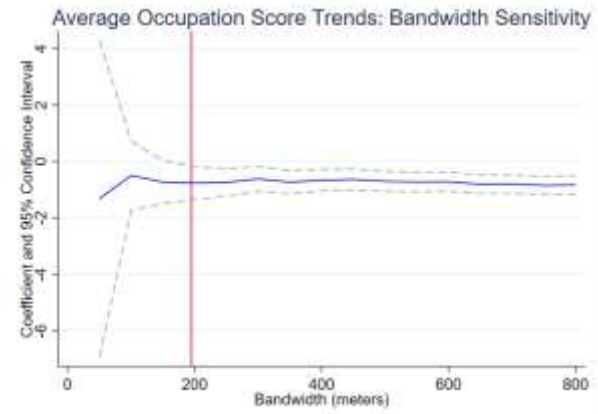
Figure A11: Optimal Bandwidth and Sensitivity of Results in Figure 3 to Different Bandwidths for the HOLC Boundary for Changes Between 1930 and 1940 in Census Measures



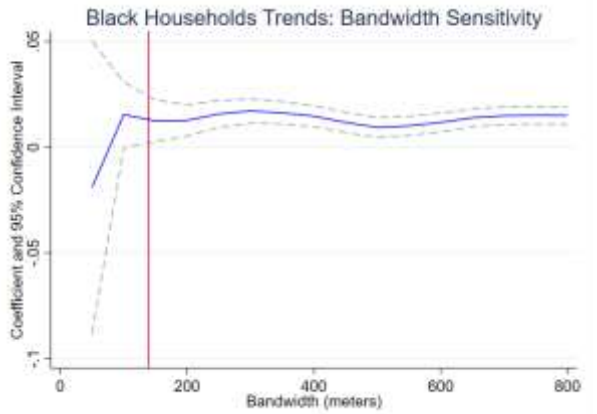
Panel A: House Values



Panel B: Rents



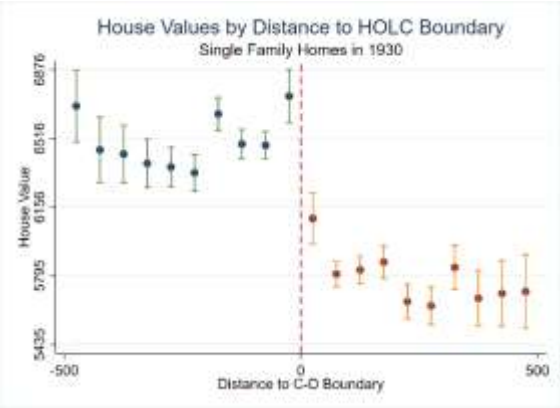
Panel C: Occupation Scores



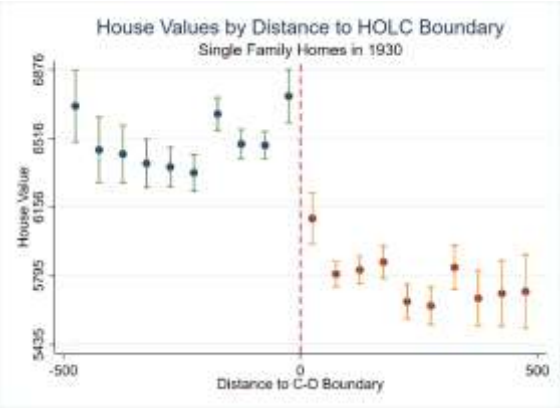
Panel D: Racial Composition

Notes: This figure graphs the estimated coefficients and 95% confidence intervals of β from equation (2) as we vary the bandwidth around a HOLC boundary. Bandwidth selection ranges from 50 to 800 meters. The red line represents the optimal bandwidth selection procedure proposed by Calonico, Cattaneo, and Titiunik (2014).

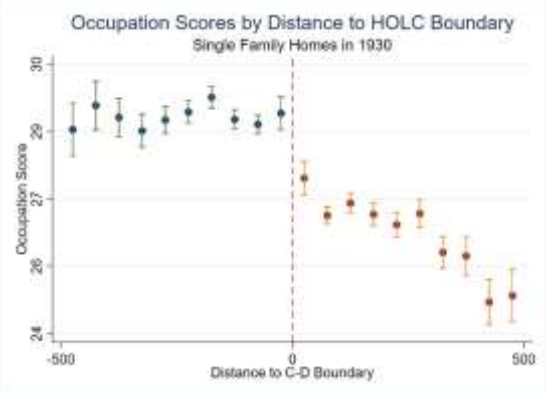
Figure A12: 1930 Levels by Distance in 50-Meter Bins to HOLC Boundary (Dropping "Thick" Borders)



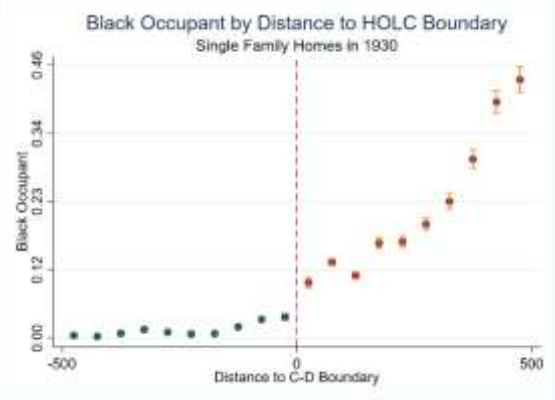
Panel A: House Values



Panel B: Rents



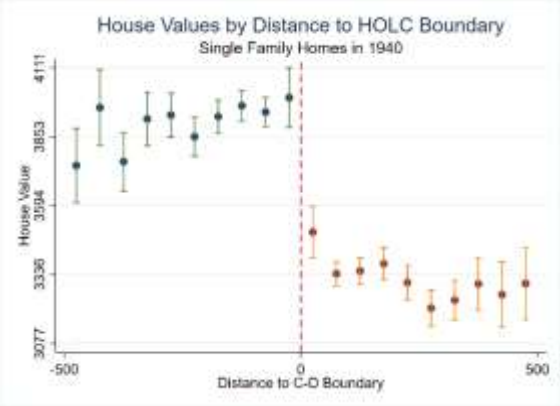
Panel C: Occupation Scores



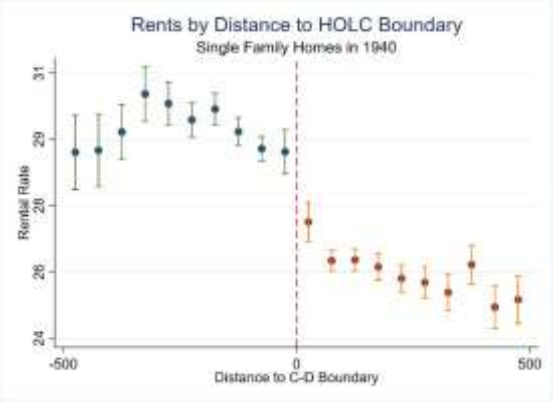
Panel D: Racial Composition

Notes: This figure shows binned averages of 1940 census data, controlling for 1930 values, for single family households by distance in 50-meter bins to a C-D HOLC boundary. The red dotted line represents the HOLC boundary, positive distances represent households in the redlined zone, and negative distances represent houses in the yellow-lined zones. All distances are measured in meters.

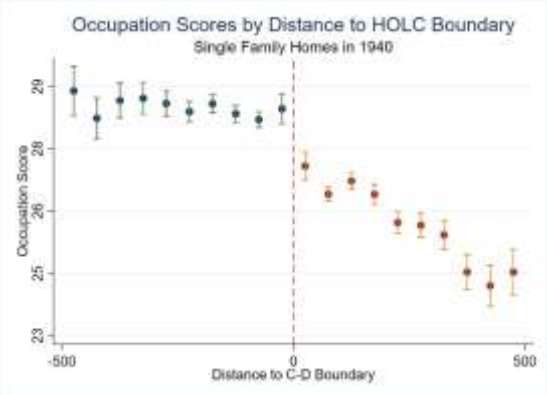
Figure A13: 1940 Levels by Distance in 50-Meter Bins to HOLC Boundary (Dropping "Thick" Borders)



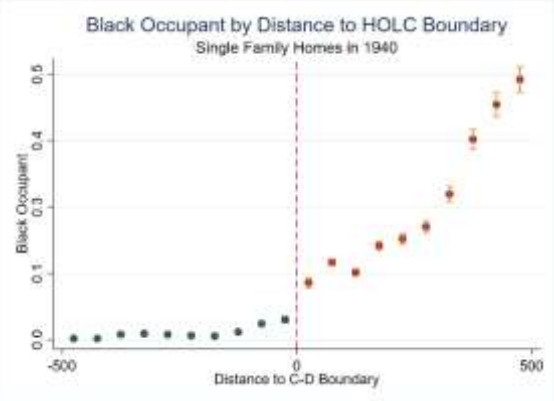
Panel A: House Values



Panel B: Rents



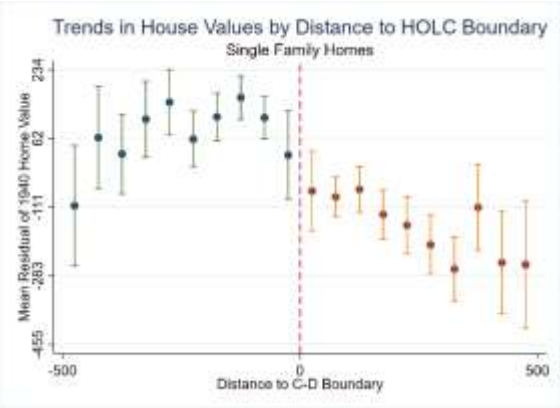
Panel C: Occupation Scores



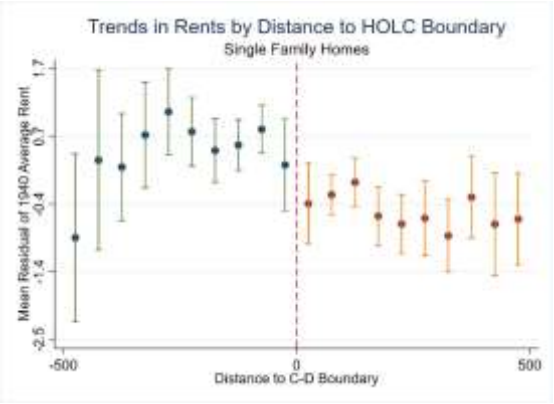
Panel D: Racial Composition

Notes: This figure shows binned averages of 1940 census data, controlling for 1930 values, for single family households by distance in 50-meter bins to a C-D HOLC boundary. The red dotted line represents the HOLC boundary, positive distances represent households in the redlined zone, and negative distances represent houses in the yellow-lined zones. All distances are measured in meters.

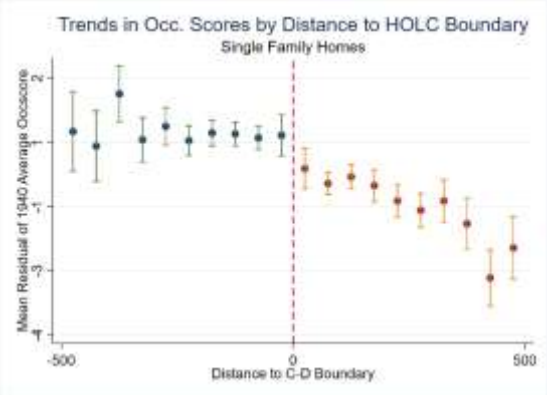
Figure A14: 1940 Trends by Distance in 50-Meter Bins to HOLC Boundary (Dropping "Thick" Borders)



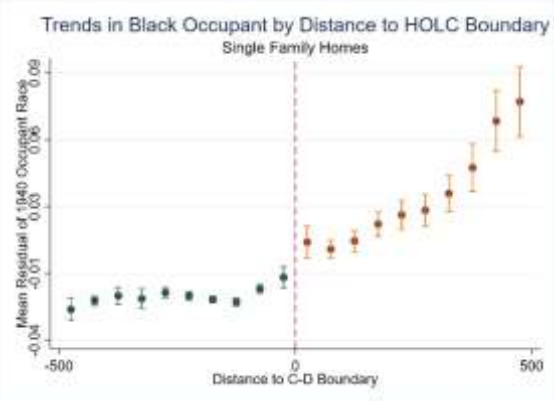
Panel A: House Values



Panel B: Rents



Panel C: Occupation Scores



Panel D: Racial Composition

Notes: This figure shows binned averages of 1940 census data, controlling for 1930 values, for single family households by distance in 50-meter bins to a C-D HOLC boundary. The red dotted line represents the HOLC boundary, positive distances represent households in the redlined zone, and negative distances represent houses in the yellow-lined zones. All distances are measured in meters.

Figure A15: Distributions of Predicted Probability the Enumeration District(ED) is Rated D Based on Other Characteristics When Percent Black Division and the Sample Coverage Are Varied

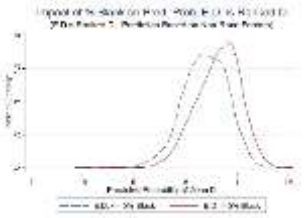
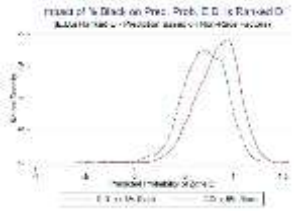
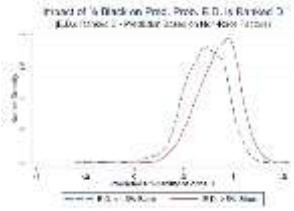
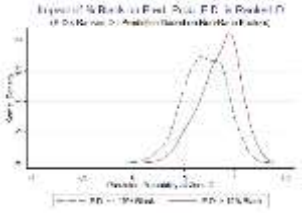
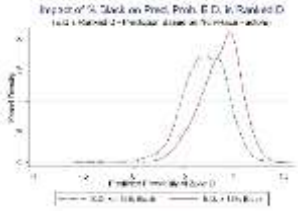
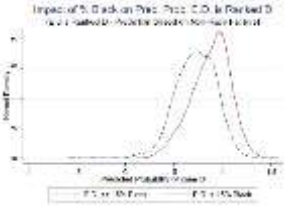
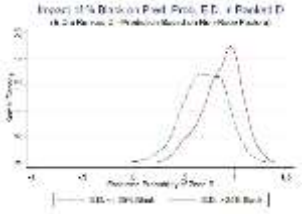
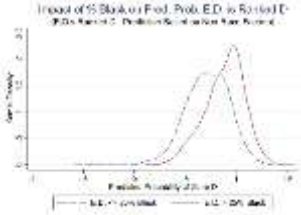
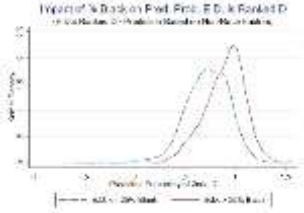
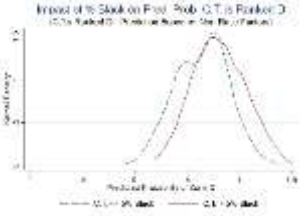
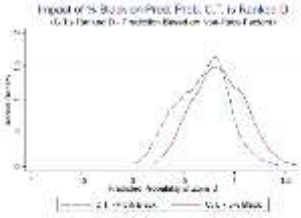
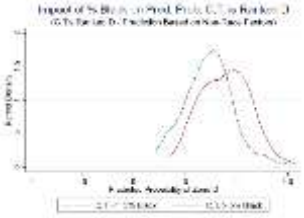
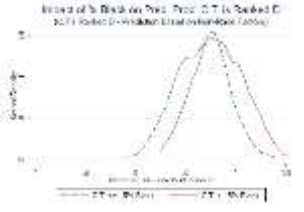
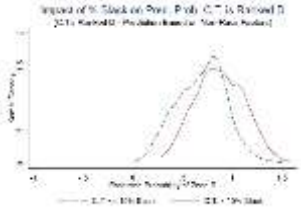
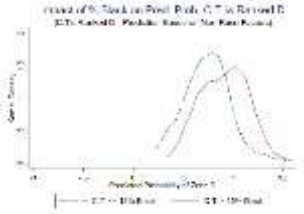
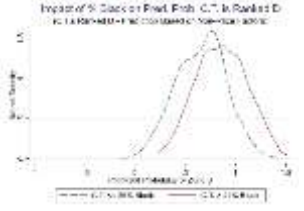
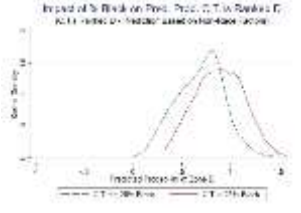
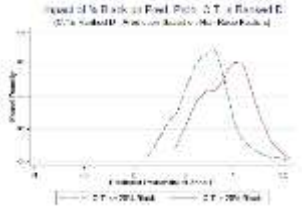
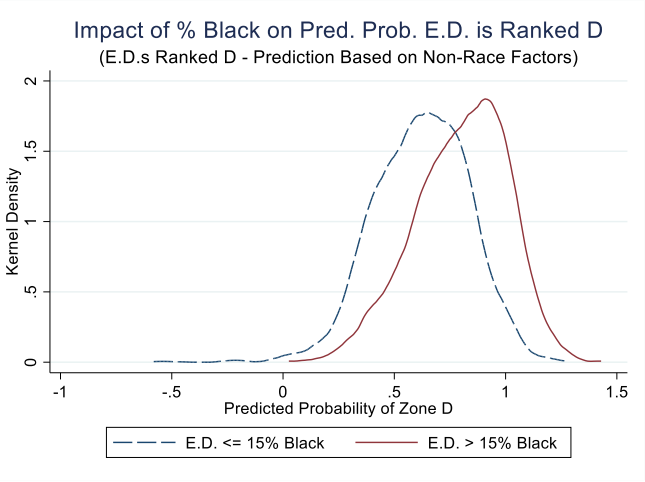
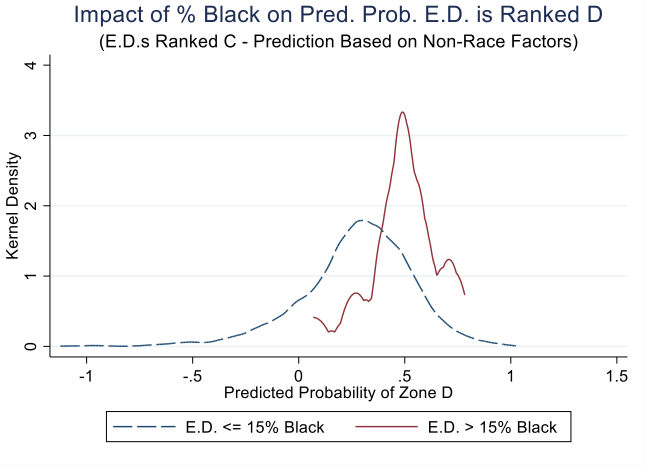
Panel A: 1930 Enumeration Districts		
 Black 5% and 50% coverage	 Black 5% and 75% coverage	 Black 5% and 99% coverage
 Black 15% and 50% coverage	 Black 15% and 75% coverage	 Black 15% and 99% coverage
 Black 25% and 50% coverage	 Black 25% and 75% coverage	 Black 25% and 99% coverage
Panel B: 1940 Census Tracts		
 Black 5% and 50% coverage	 Black 5% and 75% coverage	 Black 5% and 99% coverage
 Black 15% and 50% coverage	 Black 15% and 75% coverage	 Black 15% and 99% coverage
 Black 25% and 50% coverage	 Black 25% and 75% coverage	 Black 25% and 99% coverage

Figure A16: Distributions of Predicted Probability the Enumeration District(ED) is Rated D Based on Other Characteristics When Percent Black is above or below 15 Percent and Regression includes City Fixed Effects



Panel A: D Zones



Panel B: C Zones

Notes: This figure shows the distribution of an enumeration districts predicted probability of being redlined. Predicted probabilities were calculated using 1930 ED census data. City Fixed Effects are netted out of the predicted probability. Data is presented based on the share of black households in an enumeration district.

Table A1: HOLC Survey Summary Statistics by Zone

	(1)	(2)	(3)	(4)
	Zone A	Zone B	Zone C	Zone D
	All	All	All	All
Family Income	16760	5908	2969	1426
	(20274)	(7278)	(5492)	(771)
House Value	17474	8722	5578	3598
	(12959)	(5335)	(2492)	(2076)
Rent	69.33	51.27	36.12	20.67
	(23.93)	(16.93)	(11.53)	(8.25)
Share Black	0	0	0.004	0.168
	(0.000)	(0.001)	(0.023)	(0.288)
Share Foreign	0.02	0.07	0.24	0.44
	(0.05)	(0.14)	(0.27)	(0.29)
Share of Houses Occupied	0.84	0.89	0.93	0.82
	(0.36)	(0.28)	(0.21)	(0.34)
Share Owner Occupied	0.76	0.73	0.65	0.46
	(0.34)	(0.27)	(0.24)	(0.33)
Construction Type				
Brick	0.873	0.713	0.489	0.395
	(0.336)	(0.453)	(0.501)	(0.490)
Frame	0.073	0.240	0.495	0.586
	(0.262)	(0.428)	(0.501)	(0.494)
Other	0.055	0.046	0.016	0.020
	(0.229)	(0.209)	(0.126)	(0.139)
Repair				
Excellent	0.389	0.081	0.008	0
	(0.492)	(0.273)	(0.091)	(0.000)
Good	0.611	0.776	0.220	0.035
	(0.492)	(0.418)	(0.415)	(0.184)
Fair	0	0.138	0.747	0.490
	(0.000)	(0.346)	(0.435)	(0.501)
Poor	0	0.006	0.025	0.475
	(0.000)	(0.076)	(0.156)	(0.500)
Observations	57	187	378	268

Notes: This data was obtained from HOLC surveys.

Table A2: Address Level Census Data Summary Statistics by Zone

Census Data Summary Statistics (Address Level)

	1930				1940				% Change			
	Zone A	Zone B	Zone C	Zone D	Zone A	Zone B	Zone C	Zone D	Zone A	Zone B	Zone C	Zone D
Household Size	3.78	4.09	4.32	4.60	3.64	3.84	4.00	4.26	-0.04	-0.06	-0.07	-0.07
	(1.54)	(1.70)	(1.85)	(2.08)	(1.46)	(1.59)	(1.73)	(1.96)				
Family Size	3.62	3.92	4.10	4.21	3.47	3.69	3.82	3.96	-0.04	-0.06	-0.07	-0.06
	(1.49)	(1.67)	(1.83)	(2.07)	(1.40)	(1.56)	(1.71)	(1.95)				
Occupation Score	33.21	31.51	29.50	26.49	34.72	31.65	29.28	26.05	0.05	0.00	-0.01	-0.02
	(11.15)	(9.98)	(9.34)	(9.40)	(12.52)	(10.50)	(9.54)	(8.91)				
House Value	9417	8138	7009	5272	7597	5652	4241	3029	-0.19	-0.31	-0.39	-0.43
	(4060)	(3638)	(3340)	(2941)	(3865)	(2791)	(2164)	(1774)				
Rent	48.73	48.20	42.09	33.06	48.06	39.39	31.77	23.70	-0.01	-0.18	-0.25	-0.28
	(19.80)	(18.44)	(16.99)	(14.58)	(21.18)	(16.48)	(13.20)	(10.34)				
Share Black	0.002	0.003	0.009	0.136	0.002	0.002	0.008	0.154	0.00	-0.33	-0.11	0.13
	(0.046)	(0.051)	(0.093)	(0.343)	(0.041)	(0.046)	(0.087)	(0.362)				
Share Foreign	0.19	0.26	0.38	0.42	0.16	0.23	0.33	0.36	-0.16	-0.12	-0.13	-0.15
	(0.39)	(0.44)	(0.49)	(0.49)	(0.37)	(0.42)	(0.47)	(0.48)				
Share Owner Occupied	0.82	0.79	0.70	0.55	0.79	0.68	0.60	0.47	-0.04	-0.14	-0.14	-0.15
	(0.39)	(0.41)	(0.46)	(0.50)	(0.41)	(0.47)	(0.49)	(0.50)				
Observations	9939	137331	310241	297905	18162	157972	301542	260702				

Notes: This table includes all households from the 1930 and 1940 census.

Table A3: Results of Zone-Level Regressions of Changes Between 1930 and 1940 in Census Outcomes as a Function of HOLC Reports of Future Trends in the Desirability of the Zone

	(1) Log House Value	(2) Log House Value	(3) Log Rent	(4) Log Rent	(5) Occupation Score	(6) Occupation Score	(7) Share Black	(8) Share Black
Downward (omitted)	-	-	-	-	-	-	-	-
Slightly Downward	0.163	0.035	0.034	0.011	0.519	0.679**	0.005	0.003
	(0.108)	(0.105)	(0.028)	(0.031)	(0.336)	(0.342)	(0.016)	(0.017)
Static	0.300***	0.163**	0.097***	0.044	1.566***	1.641***	0.001	0.002
	(0.091)	(0.081)	(0.025)	(0.028)	(0.347)	(0.352)	(0.006)	(0.008)
Slightly Upward	0.361***	0.160	0.144***	0.056	2.127***	2.160***	-0.005	-0.005
	(0.083)	(0.105)	(0.031)	(0.040)	(0.458)	(0.502)	(0.006)	(0.007)
Upward	0.462***	0.305**	0.218***	0.104**	4.088***	3.777***	-0.006	-0.002
	(0.108)	(0.125)	(0.043)	(0.047)	(0.680)	(0.686)	(0.004)	(0.006)
City Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zone Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	425	425	425	425	425	425	425	425
R-squared	0.235	0.483	0.780	0.807	0.696	0.708	0.890	0.891

Notes: Robust standard errors are shown in parenthesis. *p < .10, **p < .05, ***p < .01. These results are presented graphically in Figures (7) and (A7).

Table A4: Levels and Trends at C-D Boundaries (Dropping "Thick" Borders) from 1930 and 1940 Census Addresses

	(1) 1930	(2) 1930	(3) 1940	(4) 1940	(5) Trends	(6) Trends
Panel A: House Values						
Redlined Side	-0.100***	-0.100***	-0.150***	-0.100***	-0.054**	-0.048**
	(0.013)	(0.012)	(0.017)	(0.017)	(0.023)	(0.022)
Panel B: Rents						
Redlined Side	-0.005	-0.022	-0.051***	-0.039**	-0.049*	0.023
	(0.018)	(0.016)	(0.018)	(0.017)	(0.029)	(0.030)
Panel C: Occupation Score						
Redlined Side	-1.436***	-1.342***	-1.456***	-1.387***	-0.842***	-0.806***
	(0.188)	(0.189)	(0.223)	(0.225)	(0.304)	(0.314)
Panel D: Share Black						
Redlined Side	0.053***	0.043***	0.043***	0.040***	0.014**	0.017***
	(0.008)	(0.007)	(0.011)	(0.009)	(0.005)	(0.005)
Optimal Bandwidth	Yes	Yes	Yes	Yes	Yes	Yes
Matched Sample	Yes	Yes	Yes	Yes	Yes	Yes
Boundary FE	No	Yes	No	Yes	No	Yes

Notes: Each coefficient is estimated from a separate regression. *p < .10, **p < .05, ***p < .01. We used the optimal bandwidth selection procedure proposed by Calonico, Cattaneo, and Titiunik (2014).

II. HOLC Surveys and Maps Data and Digitization

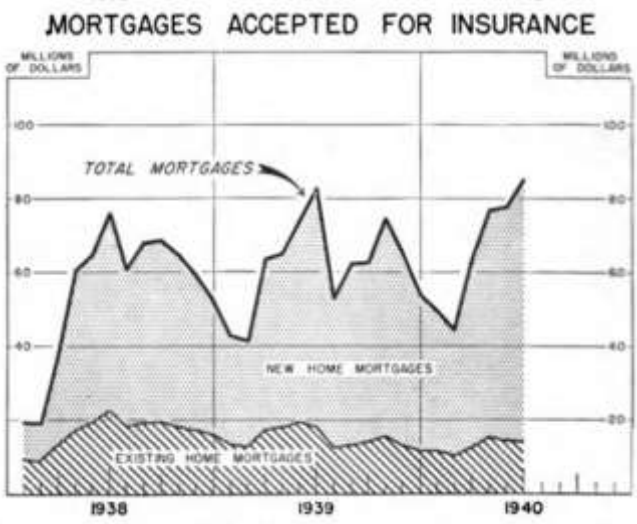
Maps and surveys were collected from the National Archives. GIS was used to georeferenced and digitize each map. Surveys collected information retroactively and dates differed for each city. Furthermore, in some instances, both high and low estimates for various types of housing (brick, frame, etc.) was reported. Housing values and rents from survey were summarized using only weighted average all reported housing types (weighted by the share of each housing type in a neighborhood) and use the first year reported from each survey.

City	Map Dates	Survey Form Dates	Price Dates Surveyed
Baltimore	May 1937	June 1937	1929, 1933-1935, 1937
Boston	February 1938	Oct 1937	1929, 1933-1936, 1937
Brooklyn	April 1938	Oct 1937	1929, 1935, 1938
Chicago	Oct 1939 - April 1940	March 1939 - Nov 1940	1935, 1937, 1939
Cleveland	March 1940	July 1939 - Oct 1939	1937, 1938, 1939
Detroit	June 1939	-	1935, 1937, 1939
Manhattan	April 1938	Oct 1937	1929, 1935, 1938
Philadelphia	June 1937	-	1929, 1933-1936, 1937
Pittsburgh	July 1937	July 1937	1929, 1933-1934, 1937
St. Louis	Oct 1940	July 1940- Oct 1940	1936, 1939, 1940

III. FHA Policies

FHA lending, while active, was relatively minor until the late 1940s, with a big increase in home lending acquiring in 1947. Home lending activity in the late 1940s was more than 5x that of the late 1930s. Furthermore, a majority of the lending that was occurring in our sample was for new home mortgages. New home lending in the 1930s was roughly 4x that of existing home lending.



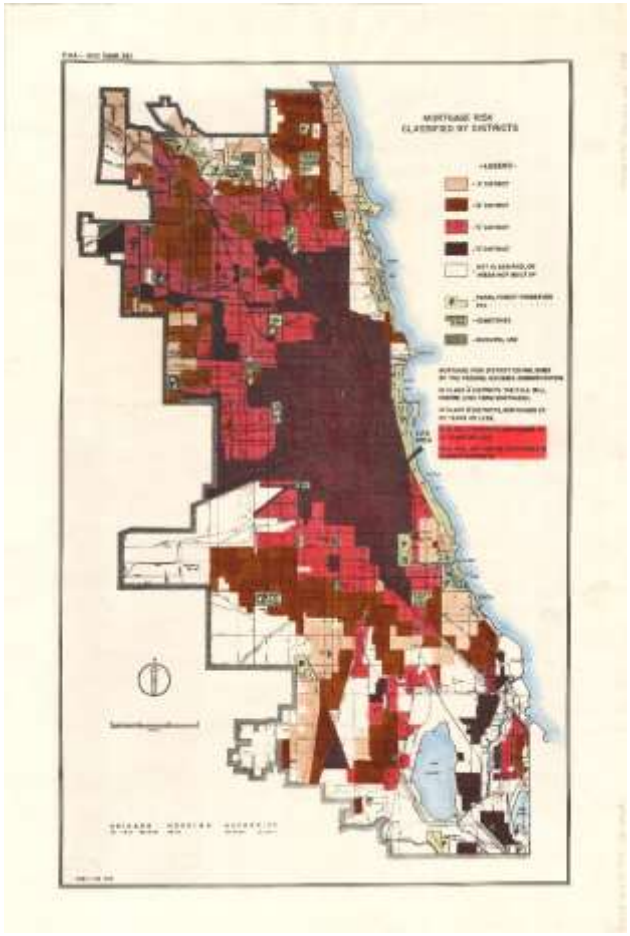


At most, 58,822 existing homes in our 10-city sample received FHA-insured mortgages between 1935 and 1940; this statistic is for "metropolitan areas" encompassing our primary city sample. Our 10-city sample contained 2,432,250 housing units in 1930. At most 2% of houses in our sample cities had received FHA-insured mortgages as of 1940 (but probably significantly less).

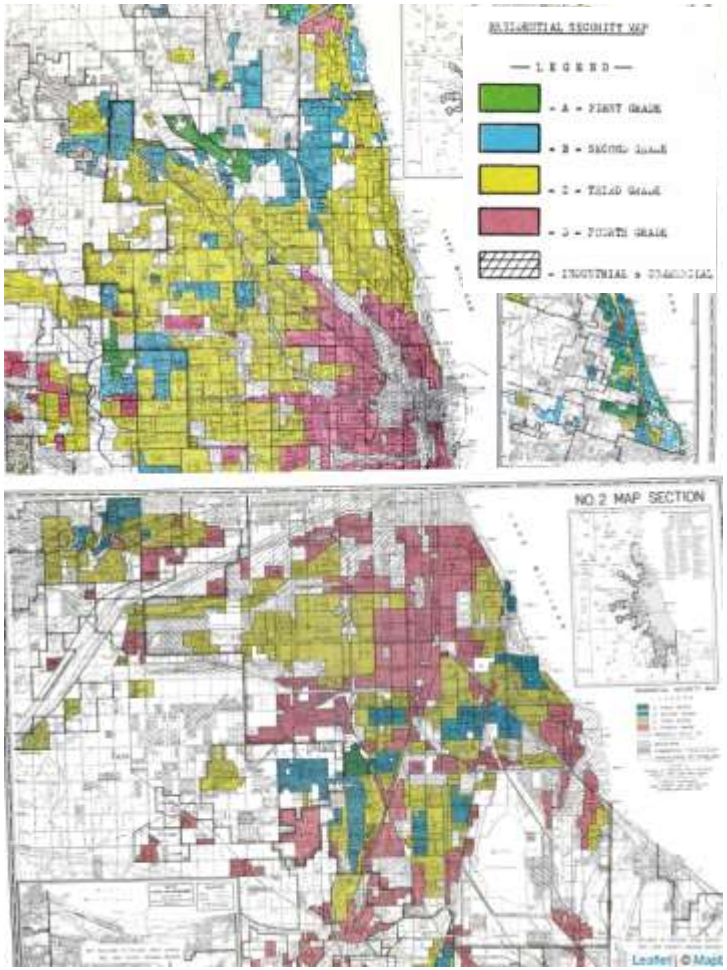
Metropolitan Area	**Mortgages 1935-1939**	**Dwellings 1930**	**Geocoded Addresses 1930**
New York	12921	557359	212225
Chicago	12373	518176	395577
Philadelphia	8258	398087	387315
Boston	1708	103141	61296
Detroit	7855	323356	194648
Pittsburgh	3350	249504	94551
St. Louis	3462	47506	111402
Cleveland	7023	209026	124151
Baltimore	1872	26095	131087
Total (sample)	58822	2432250	1712252

Lending Risk Maps for Chicago, IL

Panel A: FHA Lending Risk Map



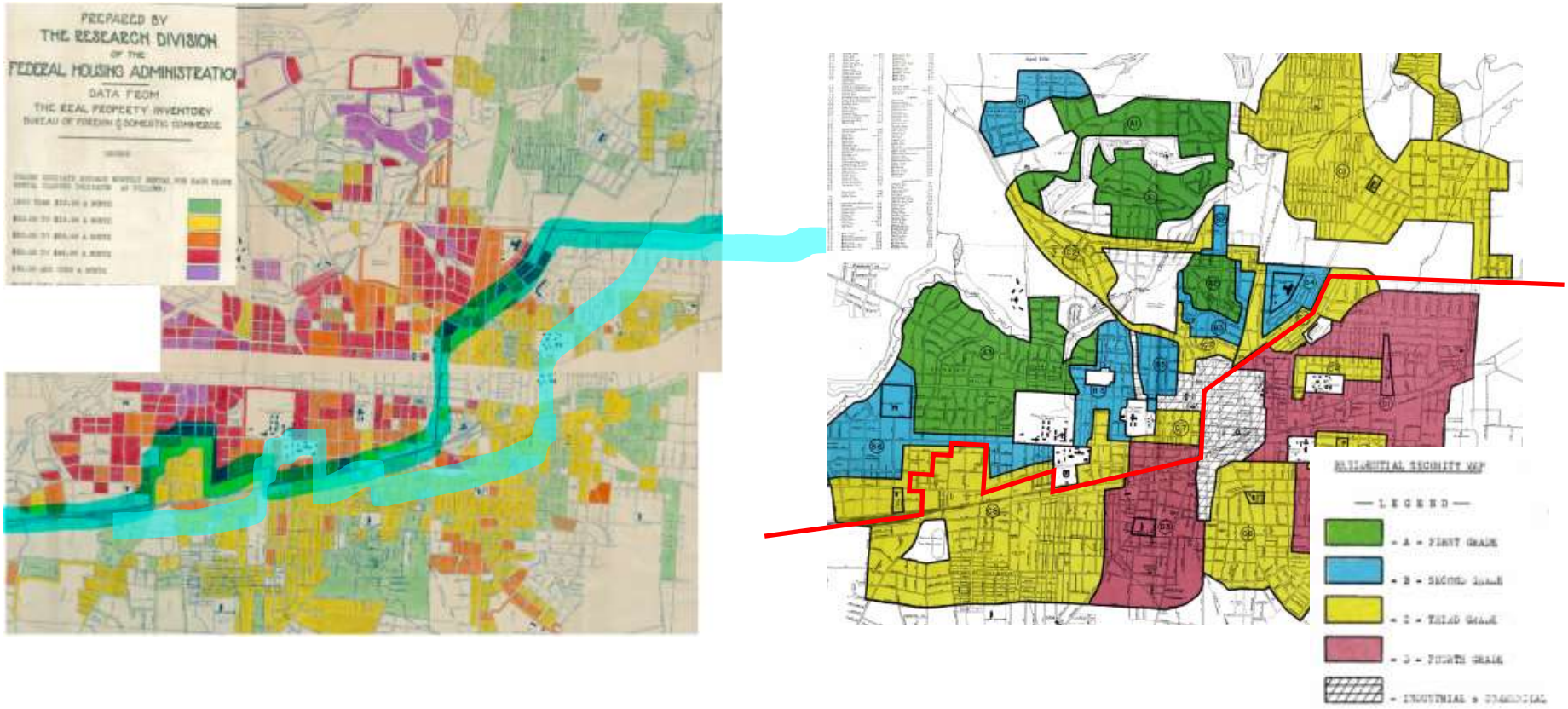
Panel B. HOLC Security Zone Map



Notes: the FHA map is from the online University of Chicago Map Collection and the HOLC map from the Mapping Inequality project at the University of Richmond.

Lending Risk Maps for Greensboro, NC

Panel A: FHA Lending Risk Map ***Panel B. HOLC Security Zone Map***



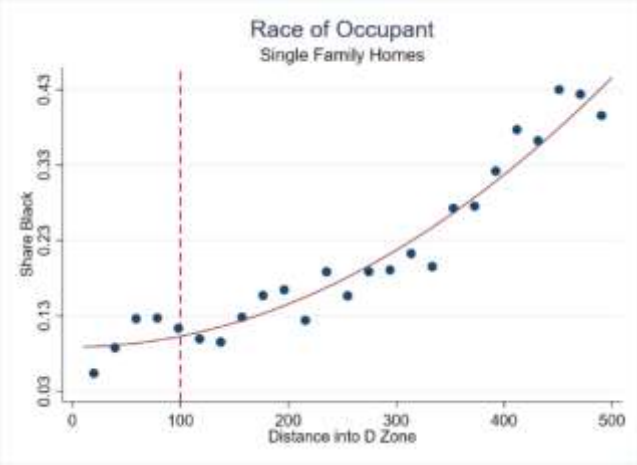
Notes: the FHA map was provided by Thomas Storrs and the HOLC map from the Mapping Inequality project at the University of Richmond. The redline added to panel B is transcribed from the actual red line from panel A (highlighted in blue) that was used to denote the boundary between high-risk and low-risk zones on the FHA map.

IV. Boundary Counterfactuals

In this appendix section we discuss the details of various counterfactual exercises used to help understand if discrimination was the driving factor in HOLC boundary placement. These counterfactuals either move black households or HOLC boundaries to explore the extent to which these changes impact the propensity of black households to be redlined. In our first counterfactual exercise we estimate the share of black residents by distance to a CD-boundary and remove all redlined black households within 100 meters of a HOLC boundary that are above the predicted value. This results in a decrease in the share of black residents who are redlined by only 3 percentage points, down from 90% to 87%.[30] Thus, while the small increase in the share of black households directly adjacent to CD boundaries could be evidence of discrimination in the delineation of HOLC zones, correcting for this bias has little impact on the overall share of black households residing in redlined areas.

We then take a more extreme approach and remove all redlined black households living within 100 meters of a HOLC boundary. Of the 13904 black households residing in single family homes in C and D zones, 12475 were redlined. Among these redlined households 2474 lived within 100 meters of the boundary. Moving these 2474 black households out of redlined areas would reduce the share of black households from 89.7% to 71.9%. This means that the great majority of black households would still reside in redlined areas. Furthermore, shifting the boundary would have shifted relative more white households into yellow-line areas and the share of redlined households occupied by black residents would have actually increased.

These exercises imply that between 80% and 97% of the redlined black households were redlined due to the overall circumstances of black families in the 1930s and thus were not the result of discrimination in HOLC maps specifically. Any discrimination in HOLC maps is dwarfed by the fact that blacks were forced to live in the lowest quality neighborhoods before the maps were created.



[30] There were 15518 households in the three bins that lie above the predicted values and within 100 meters of a CD boundary; 1924 of these households were occupied by black families. We remove 392 households to move these bins to the predicted value represented by the fitted line in the graph above.