REM WORKING PAPER SERIES

When are D-graded neighborhoods not degraded? Greening the legacy of redlining

Alba Miñano-Mañero

REM Working Paper 0353-2024

October 2024

REM – Research in Economics and Mathematics
Rua Miguel Lúpi 20,
1249-078 Lisboa,
Portugal

ISSN 2184-108X

REM – Research in Economics and Mathematics

Rua Miguel Lupi, 20
1249-078 LISBOA
Portugal

Telephone: +351 - 213 925 912
E-mail: rem@iseg.ulisboa.pt

https://rem.rc.iseg.ulisboa.pt/

https://twitter.com/ResearchRem

https://www.linkedin.com/company/researchrem/

https://www.facebook.com/researchrem/

When are D-graded neighborhoods not degraded? Greening the legacy of redlining

Alba Miñano–Mañero[*†]
ISEG, University of Lisbon
REM
UECE

February 2024

Please click here for the latest version of the paper

Abstract: This paper explores how geography shapes the legacy of redlining, the systemic mortgage lending bias against minority US neighborhoods. On average, redlined neighborhoods lag behind adjacent, less-discriminated areas in home values, income, and racial composition. Yet, redlined neighborhoods near parks and water fare better. To help understand convergence, we inventory waterfront renovations, apply machine learning to historical imagery to track tree canopy changes, and instrument such changes exploiting tree replacements due to geographic variation in tree plagues and susceptible species. Findings suggest that enhancing waterfronts and increasing tree canopy can mitigate the long-lasting effects of institutionalized discrimination.

Key words: redlining, geography, natural amenities, waterfronts, tree canopy

JEL classification: R23

[*]I am grateful to my advisor, Diego Puga, for all his guidance and mentorship. I also thank Elisabet Viladecans-Marsal and Dani Arribas-Bel for all their comments and suggestions. I have also benefited from discussions with Manuel Arellano, Dmitry Arkhangelsky, Monica Martinez-Bravo and Tom Zohar. I extend my gratitude to Francisco J. Goerlich, Camille Hémet, Jorge de la Roca, Rosa Sanchis-Guarner and Ana Moreno-Maldonado for serving in my Thesis Defense Committee. I am equally grateful for feedback from attendants to CEMFI's BUDIE and PhD Workshops, and of seminar and conferences. I gratefully acknowledge funding from Spain's Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033) through its María de Maeztu Units of Excellence program (grant CEX2020-001104-M).

[†]ISEG, Rua Miguel Lupi 20, 1200-781, Lisbon (e-mail: alba.minano@iseg.ulisboa.pt; website: https://albaminanomanero.github.io).

1. Introduction

Spatial inequalities are large and persistent. Rooted in initial conditions or geographical factors, some places consistently experience disadvantages while others perpetually thrive (Sampson, 2016; Voth, 2021). The experience of growing up in disadvantaged neighborhoods significantly shapes individuals' lives and continues to limit their opportunities as adults (Chetty, Hendren, and Katz, 2016; Sampson, 2019). Evidence regarding the efficacy of relocating individuals away from such neighborhoods is mixed and implementing such policies on a large scale can pose substantial challenges (Chyn and Katz, 2021). Therefore, it becomes necessary to think about interventions that help the convergence of persistently lagging areas.

In the case of the US, the consequences of historical policies that restricted credit in minority neighborhoods can still be felt in many of these disadvantaged areas. This paper finds a class of interventions that work by molding seemingly unmodifiable features — natural amenities — to revert that legacy. The transformation of industrial waterfronts into pedestrian-friendly promenades or the greening of sidewalks through tree-planting efforts can significantly contribute to mitigating the lasting effects of pervasive historical discrimination of minority neighborhoods.

Redlining, the historical practice of systematically denying mortgages to minority neighborhoods, is one of the main instances of institutionalized discrimination in the United States. Foreclosures became so prevalent during the Great Depression that the Federal Government began insuring mortgages through the Home Owners' Loan Corporation (HOLC) and the Federal Housing Administration (FHA). Both institutions facilitated a rapid expansion of credit and home ownership in the United States, but not among minorities, particularly African Americans. The HOLC developed an appraisal system that graded neighborhoods from A to D, outlined in maps in colors green to red. The racial composition was decisive in determining grades: D-graded (redlined) neighborhoods were those with high shares of minority residents and were systematically denied mortgage insurance.

By institutionalizing and reinforcing the discriminatory practices of realtors and lenders against minorities, redlining reduced home ownership amongst them, limiting opportunities for wealth accumulation and social and geographic mobility. Since this discriminatory practice was institutionalized at the neighborhood level, it also resulted in decades of lower property tax revenues and public and private investment. Almost half a century after the 1968 Fair Housing Act and the 1977 Community Reinvestment Act prohibited the practice, redlined neighborhoods still have lower average home values, incomes, and non-minority presence than similar nearby neighborhoods subject to weaker lending restrictions.

Prior research on the legacy of redlining focuses on its average effects (Appel and Nickerson, 2016; Rothstein, 2017; Krimmel, 2018; Aaronson, Hartley, and Mazumder, 2021; Hynsjö and Perdoni, 2023). However, I find that persistence is heterogeneous: not all D-graded neighborhoods have remained degraded, and natural amenities affect their evolution. The convergence in home values, family incomes, and non-minority shares between D-graded neighborhoods and similar (and neighboring) areas subject to less stringent historical restrictions is greater when D-graded neighborhoods feature waterfronts and parks.

And yet, geography is not necessarily destiny. In fact, I show that what helps redlined coastal and riverside neighborhoods converge is waterfront revitalization projects, for which I construct a complete inventory. Of course, waterfront revitalization is only an option by water bodies. In contrast, vegetation can be planted nearly everywhere. Identifying vegetation on a wide scale is easy today, thanks to high-resolution near-infrared imagery. To also identify trees in earlier periods, for which near-infrared imagery is unavailable, I train a machine-learning algorithm with such modern imagery to detect trees in traditional color aerial photographs. This allows me to construct the first spatially-detailed long panel of vegetation changes in US cities. Leveraging this panel, I show that greening redlined neighborhoods by planting trees has promoted convergence. A natural worry is that growing tree coverage may be a consequence of gentrification. Exotic tree plagues force neighborhoods to replace susceptible tree species with several new trees planted for every tree removed. Exploiting this exogenous source of expanded tree coverage, I establish that doubling local tree canopy is enough for redlined neighborhoods to achieve full convergence.

Multiple policies articulated the systemic discrimination towards minorities regarding their residence in the early twentieth century. Discriminatory zoning deterred the entry of minority residents into majority neighborhoods through density restrictions, and also concentrated manufacturing activity in minority neighborhoods (Shertzer, Twinam, and Walsh, 2016; Twinam, 2017, 2018). Private covenants explicitly forbade selling houses to minority households, especially African Americans (Sood and Ehrman-Solber, 2023; Almagro and Sood, 2023). While widespread, these discriminatory practices arose locally. Redlining, instead, was a nationwide practice institutionalized by Federal agencies. The discovery of the HOLC maps by Jackson (1980) was followed by city studies exploring the determinants of the assigned grades and their effects on credit access (Hillier, 2003, 2005; Crossney and Bartelt, 2005; Fishback, 2014).[1] Exploiting bordering discontinuities in the assigned grade (Appel and Nickerson, 2016; Krimmel, 2018; Aaronson, Hartley, and Mazumder, 2021) or focusing on city-level effects (Faber, 2020; Anders, 2023; Hynsjö and Perdoni, 2023), recent literature shows that redlining has persistent effects related to increased segregation and neighborhood decay.

This paper is shaped around two main conceptual contributions: heterogeneity in the persistence of redlining and the malleability of geography as a driver of such heterogeneity. However, empirically addressing those questions requires careful data treatment. For this reason, I also make two methodological contributions. The first methodological contribution is the development of a workflow to create panels of tree canopy in the presence of limited

[1] Fishback, Rose, Snowden, and Storrs (2022) note that it was mainly the Federal Housing Administration (FHA) rather than the HOLC that systematically discriminated against minority neighborhoods and started doing so before the HOLC maps were created. Thus, some scholars regard the HOLC maps (jointly produced by the HOLC and local brokers) as reflecting rather than originating the prevailing discriminatory appraisal guidelines of America at that period. Nevertheless, while the FHA's own maps were intentionally destroyed, Aaronson, Hartley, and Mazumder (2021) show an 86% overlap in areas redlined by the FHA map recovered for Chicago and the corresponding HOLC map.

training data. The second contribution is to develop a new procedure to overcome the misalignment between the HOLC maps and Census data. Since my empirical strategy implements a difference-in-difference approach between similar and adjacent neighborhoods assigned different grades, it is important not to blur the border between grades. Thus, in contrast to the rest of the literature (Appel and Nickerson, 2016; Krimmel, 2018), which apportions grades to Census units, I apportion Census data to the original graded neighborhoods.

My first conceptual contribution to recent research on the legacy of redlining is to show that persistence is heterogeneous. A broader literature on economic geography and history shows how persistent spatial inequalities often originate in historical events or policies (see Hanlon and Heblich, 2022, for a review). Geographic features have also been found to have persistent effects, primarily by being a source of location advantage that, absent large shocks, remains locked in long after that advantage ceases to be relevant (Bleakley and Lin, 2012; Michaels and Rauch, 2018). Geographic features can also act as amenities or disamenities that drive spatial sorting and inequalities (Rappaport and Sachs, 2003; Rappaport, 2007; Lee and Lin, 2018; Heblich, Trew, and Zylberberg, 2021). This paper pits the role of water and vegetation geographic features as potential amenities against the legacy of discriminatory housing policies. It shows that these amenities can be strong enough to mitigate and even eliminate the persistence of historical discriminatory policies.

Another key conceptual contribution of this paper is to show that geography is malleable. Although historical accounts suggest that the role of geography is shaped by technological advancements in industry and commerce (Jackson, 1987; Boustan, Bunten, and Hearey, 2018), most research treats location fundamentals as immutable. In contrast, I show that interventions can mold geography to create amenities where they were absent or even turn disamenities into amenities. In particular, I show that the role of waterfront locations in fostering convergence for some redlined neighborhoods is driven by waterfront revitalization projects that have turned around former industrial waterfronts. Likewise, tree coverage changes over time, and the expansion of tree canopy has helped D-graded neighborhoods overcome the legacy of redlining. This ties to a tradition of research in urban planning that studies public interventions aimed at improving neighborhoods (see Zuk, Bierbaum, Chapple, Gorska, and Loukaitou-Sideris, 2018, for a review).

Among economists, research on placed-based policies has mostly focused on infrastructure investment and enterprise zones (Neumark and Simpson, 2015), although there are a few studies on urban renewal plans (Rossi-Hansberg, Sarte, and Owens, 2010; LaVoice, 2019; Shi, Hartley, Mazumder, and Rajan, 2022). A key reason why economics research on urban renewal interventions targeting natural amenities is limited is the lack of data on relevant changes in geography. To overcome that limitation, I collect new data on waterfront improvements in abandoned industrial areas. The improvements led to the development of shoreline or riverbank parks and improved waterfront access.

A second type of intervention I consider is expansions in tree canopy, thus connecting with a broader literature on the impact of urban tree canopy. My contribution to this literature is twofold. First, I solve limitations on data availability of urban canopy by implementing a new methodology to construct panels of canopy from aerial images. This builds on work by Yang, Wu, Praun, and Ma (2009) and Bosch (2020), adding to it a method to automatize the creation of training data that is transferable across periods leveraging visual graphic techniques. This transforms a process conventionally applicable only to small areas into one that can be applied to a large set of urban areas across multiple periods. The workflow can predict the presence of tree coverage at the pixel level, even with limited training data. One of the main strengths is its potential for widespread application, given that multi-spectral aerial imagery is publicly available in multiple periods and geographic locations.

Second, I develop an instrumentation strategy to tackle endogeneity in changes in tree coverage. This connects with the extensive literature on the impact of urban trees on economic, social, and environmental outcomes.[2] Only a handful of papers attempt to estimate the causal impact of trees. For instance, Wachter and Wong (2008) does so by exploiting the design of tree plantation initiatives in Philadelphia. Particularly close to this paper are Kondo, Han, Donovan, and MacDonald (2017) and Han, Heblich, Timmins, and Zylberberg (2021), which leverage exposure to a specific tree plague in a particular city as a source of reductions in tree coverage. My instrumentation strategy exploits a different exogenous variation: the increases in tree coverage associated with replacements induced by exotic tree plagues in areas where susceptible tree species are prevalent.

The rest of the paper is organized as follows. I discuss the data construction in Section 3. This describes the Census-to-Redlining Constant Crosswalks, the new procedure to match the Census data and the HOLC maps. It then introduces the workflow used to predict tree canopy in aerial images and the construction of the waterfront modifications data. I also describe how to leverage the data on the presence of exotic tree plagues and their potential tree hosts to construct an instrument for changes in the tree canopy.

Section 4 estimates the average persistence of redlining on home values, family income, and neighborhood racial composition. To do so, I estimate a difference-in-difference that compares neighborhoods subject to the most severe credit restrictions (D-graded or redlined) with nearby areas experiencing less stringent policies (C-graded) before and after legislation prohibiting redlining practices. The nearby D-C comparison minimizes endogeneity concerns as surveyors considered them the most similar, and being nearby, they are also likely to share the same local unobservables. The findings of Section 4, leaving aside methodological differences, align with the recent literature on the persistence of redlining: decades after the outlawing of redlining policies, D-graded neighborhoods still have a less white population, home values, and income.

The first significant departure relative to prior research is in Section 5, where I show that redlining's negative legacy is highly heterogeneous. By leveraging data on the location of

[2]For instance Pandit, Polyakov, and Sadler (2014), Morales (1980), Netusil, Chattopadhyay, and Kovacs (2010) and Franco and Macdonald (2018) on housing prices; Holtan, Dieterlen, and Sullivan (2015) on social capital, Hoffman, Shandas, and Pendleton (2020) on redlining and urban heat island effects or Kondo, Han, Donovan, and MacDonald (2017) on crime.

the shoreline, lakes, rivers, and parks, the results show greater convergence in demographic composition and housing values for D-graded neighborhoods proximate to water and green amenities.

Section 6, however, shows that what drives faster convergence in waterfront areas is not water per se but interventions that turn it into an amenity. Neighborhoods that experience reduced persistence are those whose waterfronts were improved and made accessible. The D-C gaps for areas with modified waterfronts get reduced by around 70% -80% in all variables. However, unmodified water features do not affect convergence significantly.

Section 7 shifts the focus to increased vegetation as a source of convergence. Tree planting has the advantage of being a widespread intervention, yet it presents other challenges: complex data construction and endogeneity concerns. I address these challenges using my novel panel of tree canopy changes and instrumenting increases in tree canopy with tree replacements in multiples of cut-down trees induced by exotic tree plagues in areas where susceptible tree species are prevalent. My results demonstrate that increased vegetation allows D-graded neighborhoods to converge. The paper concludes with Section 8.

2. The historical context of redlining

The Civil Rights Act of 1866 codified equal rights for all races, including regarding home ownership. In 1917, in *Buchanan v. Warley*, the Supreme Court forbade local ordinances that explicitly segregated population. Nevertheless, discriminatory and segregationist practices within housing markets remained in place for much longer. These practices against minorities operated through subtler means that circumvented these legal prohibitions, with redlining as a prominent example.

Redlining arose during the housing crisis that followed the Great Depression. A typical house valued at $5,000 in 1926 was only worth $3,300 in 1932, while home foreclosures rose from 68,000 in 1926 to 250,000 in 1932 (Jackson, 1987). In 1933, more than 1,000 loans were foreclosed daily, and half of the home mortgages were in technical default (Jackson, 1987; Wheelock *et al.*, 2008). The annual foreclosure rate continued to exceed 1% until 1935 and only returned to 1926 levels by 1941 (Wheelock *et al.*, 2008).

The administration initiated a series of reforms to stabilize the housing and mortgage markets and assist distressed borrowers. The first attempt, the Federal Home Loan Bank Act, arrived in July 1932 and established a system of Federal banks to act as discount banks for home mortgages with a corresponding supervision system (the Federal Home Loan Bank Board, FHLBB). However, of the 41,000 homeowners who directly applied for loans during the first two years of the Act, only three were approved (Jackson, 1987). Effective housing measures only started to be implemented after President Roosevelt took office in 1933.

Roosevelt's New Deal administration created new institutions to intervene in the housing and mortgage markets. The Home Owners' Loan Corporation (HOLC) was established in 1933 and started to operate as part of the FHLBB to substitute for the inefficient loan provision of the Federal Home Loan Bank Act. Initially, it acted as a "bad bank" issuing bonds to buy mortgages from distressed borrowers and provide them with better conditions.

The HOLC was conceived as a temporary emergency actor in charge of assisting borrowers who could not access private refinancing mortgage markets. However, the magnitude of the foreclosure crisis led to a sizeable intervention. Between 1933 and 1936, the HOLC provided one million low-interest, self-amortizing, long-term, and uniform-payments mortgages. These mortgages amounted to a total of over $3 billion and one out of five dwellings received HOLC financing (Harriss, 1951; Hillier, 2003).

The scale of refinancing by the HOLC triggered concerns that mortgages could go foreclosed even after refinancing, leaving the Government with assets whose value was unknown. In fact, over the existence of the HOLC, 19% of its loans were foreclosed, with foreclosure rates being as high as 40% in New York, New Jersey, and Massachusetts (Harriss, 1951). By mid-1935, with one-third of the eventually foreclosed HOLC loans being already delinquent for several months, the FHLBB established the City Survey Program, shifting the primary focus of the HOLC to this new initiative. The goal was to develop a standardized system to assess the value of the real estate now owned by the Government while ensuring the stability of the mortgage market.

With the establishment of the City Survey Program, the HOLC introduced a systematic property appraisal process based on neighborhood characteristics.[3] The surveys did not aim to guide the HOLC refinancing, which was already almost complete, but rather to help manage the portfolio of HOLC assets and guide the sale of the foreclosed properties (Fishback, Rose, Snowden, and Storrs, 2022). Between 1935 and 1940, the HOLC evaluated neighborhoods in the 239 cities with a population greater than 40,000 inhabitants. The appraisal process lead to the creation of the *Residential Security Maps*, commonly known as the redlining maps, due to the ink used to color the neighborhoods deemed riskiest for lending purposes.

The HOLC surveyors worked with local appraisers and lenders to create the redlining maps (Hillier, 2003; Winling and Michney, 2021). Following the FHLBB appraisal manual, neighborhoods were classified into four categories reflecting the desirability of lending in the area. These four categories were assigned the grades A, B, C, and D, from most to least desirable, and were colored green, blue, yellow, and red on the maps. More specifically, the FHLBB Appraisal Manual described the grades as follows (Hillier, 2005):

- A-graded, greenlined: "Best" neighborhoods were "homogeneous in demand in good and bad times."
- B-graded, bluelined: "Still Desirable", "like a 1935 good automobile, but not what people who can afford it are buying today."
- C-graded, yellowlined: "Definitely declining" neighborhoods that were "suffering from an infiltration of lower grade population."

[3]Appraisals were common before the HOLC started to conduct them. The relevance of the HOLC appraisal system was "the creation of a formal and uniform system of appraisal, reduced to writing, structured in defined procedures, and implemented by individuals only after intensive training. The ultimate aim was that one appraiser's judgment of value would have meaning to an investor located somewhere else." (Jackson, 1987, p. 197).

- D-graded, redlined: “Hazardous” neighborhoods where “the things that are now taking place in C have already happened.”

As evidenced by the area description files, the appraisal process reflected the institutionalized racism of the period resulting in the systematical undervaluation of black, immigrant, Jewish, or racially mixed neighborhoods (Jackson, 1980; Hillier, 2003).[4] For instance, in C- and D- graded areas, there were “heavy concentrations of low grade aliens” as in Detroit, or in Staten Island where “Italian infiltration depress residential desirability in this area.” “Slow increases of subversive races” were taking place in Los Angeles and “coloured infiltration” was “a definitely adverse influence on neighborhood desirability” in Brooklyn. Areas with a “community of the best class of Negroes” as the historical upper-class black communities of Jacksonville were also redlined.

According to Jackson (1987), the appraisal process was based on the prevalent ecological and socioeconomic theory of neighborhood change at the time. Appraisers believed that the racial composition of the neighborhood determined the housing value.[5] They also saw neighborhood decline as inevitable due to the increasing age and obsolescence of housing and the consequent filtering towards lower-income groups. As a result, black and minority neighborhoods would receive unambiguously the worst grades. Neighborhoods with low rents and aging housing prone to filtering down soon would be in the second worst grade. The best grades were reserved for the newer parts of the city and for areas that could protect from the “infiltration” of population groups that represented “adverse influences” for housing values stability (Hillier, 2003) through zoning restrictions or private covenants.[6] Although D-graded and C-graded areas shared similarities in neighborhood demographics and housing characteristics, D-graded areas were the ones considered to constitute a lending risk for banks, and the recommendation was that credit should be restricted or avoided.

While it is unclear how publicly available the HOLC maps were, the HOLC is regarded as the primary actor behind the institutionalization of redlining due to the development of its standardized appraisal process.[7] The active refinancing program of the HOLC ended in 1936, before the City Survey Program began. However, the collaboration of HOLC agents and local brokers contributed to the homogenization of appraisal criteria, implying that active lenders followed similar grading techniques (Winling and Michney, 2021).

The influence of neighborhood characteristics in appraisals was also shared by the Federal Housing Administration (FHA), created by the National Housing Act of 1934. Differently from the HOLC, the FHA was designed as a long-term agency to reform and stabilize the

[4]The area description files are available together with the redlining maps on the Mapping Inequality Project of the University of Richmond (Nelson, Winling, Marciano, Connolly *et al.*, 2017).

[5]The FHA (1936) appraisal manual mentioned that “the infiltration of inharmonious racial groups [...] tends to lower the levels of land values and to lessen the desirability of residential areas.” (FHA, 1936, p. 72).

[6]This has been corroborated by Hillier (2005), Fishback (2014) and Crossney and Bartelt (2005) since they show that both the racial composition and housing characteristics were determinants of the grades in the particular cities they study.

[7]Researchers like Hillier (2003) and Greer (2013) maintain that the maps were not diffused despite the high demand for them, while others like Jackson (1980) and Woods (2012) defend the opposite.

mortgage sector. It had two main goals: substitute for the collapsed private guaranty sector by offering public insurance to private mortgages and incentivize residential construction by directing attractive insured loans to new developments. By the late 1940s, the FHA was providing insurance for one-third of the new homes (Aaronson, Hartley, and Mazumder, 2021), and by 1972, the FHA had insured mortgages for eleven million families (Jackson, 1987). The FHA contributed to the decay of core areas through its predilection towards single-unit rather than multi-unit housing, by offering worse conditions for repair loans, and by virtually only allowing insurance in suburban areas through its lending guidelines and construction standards. Moreover, the National Housing Act established that only "economically solid" projects could be insured, increasing the FHA concern about neighborhood risk. As a result, neighborhood risk ratings were employed from the onset of the FHA (Fishback, Rose, Snowden, and Storrs, 2022). This excluded minority neighborhoods and populations from mortgage insurance.

Similarly to the *Residential Security Maps*, the FHA created its own lending risk maps. Given the simultaneity between the FHA ratings and its insurance activities, research on redlining would have ideally focused on this agency. However, the FHA destroyed the maps when facing lawsuits for discrimination. The justification to use the HOLC maps to study redlining is rooted in the prevalent view that HOLC appraisal guidelines determined the FHA ones (Hillier, 2003). Recent research indicates that the FHA had access to the HOLC maps and that there was constant communication between both agencies (Michney, 2022). Also, the surviving FHA map of Chicago is remarkably similar to the HOLC one (Aaronson, Hartley, and Mazumder, 2021).

One may never be able to positively assert whether the HOLC maps were publicly used or not. Nonetheless, by reflecting the prevalent appraisal guidelines of America at that period, including those present in the destroyed FHA maps, the HOLC maps serve as an approximation to the discriminatory lending practices of the time. Nevertheless, our results should not be interpreted as an outcome solely attributable to the specific maps or the HOLC, but rather as the consequence of the consistent and homogenized historic practice of redlining (Fishback, LaVoice, Shertzer, and Walsh, Forthcoming).

The outlawing of redlining practices was a gradual process that started with the passing of the 1968 Fair Housing Act, which prohibited all kinds of discrimination in housing markets. However, community groups continued denouncing widespread wrongful credit denials in minority neighborhoods. Tabulated mortgage data from the 1975 Home Mortgage Disclosure Act (HMDA) allowed these groups as well as Congress to substantiate ongoing housing discrimination. This lead to the passing of the Community Reinvestment Act (CRA) in 1977. This effectively outlawed discriminatory lending based on neighborhood characteristics by establishing that banks should assess and meet the financial needs of the low and moderate-income neighborhoods of the communities they served. For this reason, in our empirical analysis we focus on the 1977 passing of the CRA as the key before-and-after event for our difference-in-difference strategy. Nevertheless, the gradual nature of the legislation process implies that neighborhoods would have began to gradually change a few years prior. And yet, despite its legal prohibition, the effects of redlining may endure due to its lasting impact on segregation, disinvestment, and wealth inequalities. I now turn to describe the data that will allow me to examine this persistence and how it varies with certain geographic amenities.

3. Data

This paper makes three data contributions. First, it leverages a new procedure to match Census data with the HOLC maps without blurring spatial discontinuities in grade assignments. Second, to assess changes in the geographical amenities of neighborhoods, it constructs a new dataset that dates and geolocates waterfront renovations. Lastly, it implements a new methodology exploiting machine learning and image segmentation to obtain panels of tree coverage from aerial imagery. The study area is neighborhoods graded by the HOLC with Census data for at least 80% of its area at the tract level in 1940.

Census-to-Redlining Constant Crosswalks

The *Mapping Inequality* project of the Digital Scholarship Lab of the University of Richmond has digitized the HOLC maps from the National Archives (Nelson, Winling, Marciano, Connolly *et al.*, 2017).[8] The result of the digitization is a collection of georeferenced maps that show the location and shape of the neighborhoods delineated by HOLC surveyors. Accompanying these maps are the grades assigned (A-B-C-D) and, if available, the area description files detailing the surveyors' rationale for these grades. For estimation purposes, neighborhoods are matched to their corresponding 2010 MSA and Census division.[9] Appendix Table 8.3.1 shows the HOLC cities considered, their corresponding 2010 MSA, and the number of neighborhoods with Census data for the 1940-2015 period. Appendix Table 8.3.2 shows the city-grade HOLC neighborhoods distributions.

To explore the effects of redlining in neighborhoods HOLC maps are matched with Census tract (1940-1980) and block-group level data (1990-2015) from the National Historical Geographic System (NHGIS). The analysis is restricted to this period because tract-level data is only available from 1930 but with limited city coverage. Hence, setting 1940 as the initial period allows for observing more cities.[10] I address the misalignment between Census data and the HOLC maps by constructing data at the HOLC neighborhood maps level with the use of a new set of crosswalks, the Census-to-Redlining Constant Crosswalks. Data at the neighborhood level is the weighted sum of the Census units data that compose the HOLC neighborhood, with

[8]When this paper was initiated (2018-2019), the count of digitized maps was slightly lower, leading to the omission of some recent additions to the *Mapping Inequality* project.

[9]The 2010 definition is used for practical purposes since it is the definition that contains most of the graded neighborhoods. The assignment is based on the largest spatial overlap.

[10]Data availability imposes the additional restriction that I cannot explore the effects of the introduction of redlining and reduces the possibility of exploring pretrends to the set of cities that were surveyed by the census in 1930 and 1940. Krimmel (2018) performs this comparison and shows no different pretrends between neighbouring D-C areas.

weights equal to the area share of the Census unit that falls within the neighborhood and belonged to a 1940 tract. To ensure neighborhoods are captured comprehensively since 1940, the procedure imposes the additional restriction that at least 80% of the neighborhood had to be covered by tracts in 1940.[11]

In contrast to assigning HOLC grades to Census units as in Hillier (2005), Appel and Nickerson (2016) and Krimmel (2018) among others, my data construction process preserves the original and sharp variation in assigned grades. It ensures a gradual change in the characteristics of adjacent neighborhoods and eliminates the measurement error caused by grade assignments. Hence, these crosswalks align with the requirements of the empirical strategy. The only arising concern would be splitting a very heterogeneous Census unit into different grades or if a graded neighborhood is composed of heterogeneous tracts. By the design of the data sources, this is a minor concern since both Census and HOLC units were drawn to capture homogeneous areas.[12]

Providing evidence on the relevance of the data construction process requires observing reliable data at the HOLC level and at the Census unit level to be able to crosswalk it to the graded neighborhood or assign to the unit the HOLC grade with the largest spatial overlap. Due to the lack of data at the HOLC level, I take advantage of high-resolution ($100m^2$ pixels) gridded population data. This data allows obtaining, with the same source and procedure, a reference-true value for the original neighborhood and Census units. To assess the impact of the data construction, I then estimate, with these three different datasets for the 2010 population, a regression of population counts and density on a D-graded variable and border-pair fixed effects. Appendix Tables 8.3.13 and 8.3.14 show that using Census-to-Redlining Crosswalks always yields the closest estimates to the true one, and differences can be substantial, as in the magnitudes for population counts.

Geographical amenities

Next, I use data for water and parks as natural amenities. The choice is motivated by the evidence showing their relevance for neighborhood outcomes and by the fact they are the amenities with enough variation among nearby areas (Jackson, 1987; Brueckner, Thisse, and Zenou, 1999; Rappaport and Sachs, 2003; Lee and Lin, 2018), or the survey on the impact of parks by Crompton and Nicholls (2019)). Data on water features is collected from the Coastal Geospatial Data project of the National Oceanic Atmospheric Administration (NOAA) and includes the shoreline, Great Lakes, any other lake, and major rivers. For parks, the data relies on the ESRI layer on parks. To capture meaningful natural amenities data for lakes and parks is restricted to the set of lakes named "lake" or "pond" and to parks containing "parks", "gardens" or "forests" as part of the name.

[11]See Appendix 8.1 for additional details on the Crosswalk construction.

[12]There is evidence of heterogeneity within neighborhoods in the area description files. However, entropy indices (not shown) in both 1940 and 2015 were, on average, around zero, meaning that the Census units in neighborhoods have essentially a very similar composition in terms of population, home values, and family income.

A neighborhood is defined as *having* water and parks natural amenities when at least 20% of its area falls within a 500-meter buffer around any of the features.[13] The area threshold was determined by visual inspection. Low thresholds do not capture meaningful situations, whereas excessively high thresholds select very specific neighborhoods. The 20% criterion balances both: it is stringent enough to capture the substantial presence of amenities and differences among neighboring areas, yet not so stringent as to raise concerns about sample selection.

Waterfront modifications

Also, I hand-collect and geolocate data on waterfront modifications in the cities under study. This dataset was created using data from a variety of sources, including departments of parks, local history and news, tourism offices, and redevelopment and planning agencies. In most cases, the redevelopment plans resulted in new parks, greenways, or promenades that can be easily geolocated. In other cases, the project districts or the coordinates of the created place serve as geolocation.[14] Neighborhoods with a modified waterfront are the ones that intersect a 500-meter buffer around a geolocated modification. Appendix Table 8.3.3 contains the list of the improvements. A detailed description of the data is available in Appendix 8.2.

Tree canopy

Typically, research exploring the role of trees has relied on tree surveys with coverage restricted to particular cities and, in very few cases, a panel dimension. Moreover, recent machine learning algorithms require training data whose availability at high-resolution and large scales is a recent phenomenon (near-infrared light, NIR) or, due to its costs, its geographic and time availability is restricted (Lidar). To overcome this limitation, I propose a new method to train data from older periods with recent NIR data and produce the first panel of tree coverage in more than 30 US metropolitan areas.

This paper implements the pixel classification algorithms developed by Yang, Wu, Praun, and Ma (2009) and Bosch (2020) on the National Agricultural Imagery Product (NAIP) to construct data on the tree coverage. The NAIP is a program developed by the US Department of Agriculture since 2003. It acquires and publishes high-resolution ($1m^2$ or less) aerial images taken during the agricultural growing season every three years since 2009. The images contain, for every $1m^2$ pixel, the red-green-blue (RGB) channels of the underlying color and, for recent years, also the non-visible NAIP band. Due to the time cost of predicting tree canopy, I limit the analysis to two periods and maximize the temporal interval, which ensures observing

[13] Distances are avoided due to the irregular shapes of graded neighborhoods. The placement of centroids, as averages of vertices, may not necessarily lie within the neighborhood boundaries, thus failing to capture the actual presence of amenities within the neighborhood.

[14] These modifications are restricted to those that were direct attempts by cities, which means that waterfronts that might have revitalized from the unplanned action of individuals by setting commercial or leisure venues are not considered.

meaningful tree canopy changes. Given that the first available year differs across states, the first period ranges between 2003/2007, and the second one between 2014/2015. Appendix Table 8.3.4 shows the periods for every city considered.

In contrast to most tree detection algorithms that are intensive in data requirements, Yang, Wu, Praun, and Ma (2009)'s method has the advantage of achieving similarly good results using only RGB data. The prediction accuracy relies on training the algorithm with precise ground-truth masks. One of the methods that is used to produce training data leverages using limitedly available NIR, which captures alive vegetation due to the reflectance properties of photosynthesis. To overcome this limitation, I employ various visual graphic techniques using modern NIR to train models that can predict periods without this light. As Yang, Wu, Praun, and Ma (2009)'s algorithm relies exclusively on RGB colors, I avoid potential inaccuracies caused by different colors across periods by equalizing the lightness and color histogram of all first-period images to their counterpart in the second period – the one with NIR data and hence used as training – as a pre-step. To find the tree pixels in the training images, I first compute the widely used normalized difference vegetation index (NDVI) as $\frac{NIR-R}{NIR+R}$. The NDVI ranges from -1 to 1, with higher values representing the densest and most alive vegetation. To account for the sensitivity of the NDVI to local and vegetation conditions, the threshold that separates not-tree and tree pixels in the training data is determined by finding the two NDVI values that maximize the variance between three-pixel classes and minimize the within-class variance (i.e., Otsu's thresholding). Since most urban areas exhibit mixed features characterized, double segmentation guarantees the highest threshold captures the class with the most alive (i.e., higher chlorophyll content) and dense vegetation, which corresponds to trees.[15]

Exposure to exotic tree plagues

I construct the change in exposure to plagues by merging the data on county presence of plagues as of December 2015 compiled by Fei, Morin, Oswalt, and Liebhold (2019) and hosts potential distribution of Wilson, Lister, Riemann, and Griffith (2013). The data on the first detection is supplemented with data from multiple sources to obtain the most accurate detection date possible and at the highest geographic resolution. The host species distribution is a raster for each tree specie, in which each $250 \times 250m$ pixel represents the predicted live-tree basal area of that specie using reference data between 2000-2009. Of the total 162 potential host species of the plagues, Wilson, Lister, Riemann, and Griffith (2013) provide the species distribution of 130. Potential host exposure in a neighborhood is computed as the ratio of the total basal area of potential hosts of a particular plague to detected tree pixels in 2000. Appendix Figure 8.3.1 shows the county distribution of the total number of selected plagues and Appendix Figure 8.3.2 illustrates the distribution of potential hosts of one of these plagues, the Emerald Ash Borer, in the city of Chicago.

[15]For further details, see the original paper by Yang, Wu, Praun, and Ma (2009) and the implementation developed by Bosch (2020). In Miñano-Mañero (2023), I describe in detail the relevance of the methodology.

Sample and variables of interest

The complete sample consists of 3,779 graded neighborhoods per decade, with approximately 62% having natural amenities. Of those, 7% have a modified waterfront. In terms of population, the data accounts for nearly 19% of the US population in 1940. However, the population is not evenly distributed among categories: despite accounting for 66% of graded neighborhoods, D and C areas contain over 81% of the sample population.There are also racial disparities in population distribution: while 97% of the black population in the sample concentrates in D and C areas, only 3% is in the best two categories. The population distribution corroborates the fact that redlining mainly affected black communities.[16]

This paper focuses on the evolution of the white population share, home values, and family income because. As discussed in Section 2 these variables determined the assigned grade, they are more likely to have been influenced by redlining, and natural amenities can affect their evolution (Villarreal (2014), Lee and Lin (2018), Heblich, Trew, and Zylberberg (2021)). Section 2 already addressed the racial aspect underlying redlining.[17] Next, I focus on home values measured as the percentage of owner-occupied housing units that are on and above the MSA median home values.[18] Because housing accounts for a large portion of household wealth, the persistent wealth gap between black and white populations may be related to the impact of redlining on segregation and depressed home values. Finally, family income is measured as the percentage of families that are on and above the MSA median family income.[19]

4. The legacy of redlining

The primary identification challenge to estimating the persistence of redlining stems from the design of the HOLC grading process. As discussed in Section 2, the assigned level of credit restrictions reflected neighborhoods' housing and demographic characteristics by the late 1930s. Because neighborhoods were already different when redlining was introduced, a traditional differencing approach will not be able to distinguish if D-graded neighborhoods evolve differently from the rest because they were on different paths to start with or because they were redlined. Appendix Tables 8.3.8 and 8.3.10 show that the discontinuities in population, housing values, and income for each grade in 1940 continued to persist by 2015. Appendix Table 8.3.11 also highlights that D-graded neighborhoods in 1950 had the higher share of their

[16]See Appendix Tables 8.3.5, 8.3.6 and 8.3.7.

[17]I follow only the white population because the population from other races affected by redlining, besides white and black populations, is negligible and concentrated in particular areas. Thus, the key differences in terms of population are between the black and white populations.

[18]For each decade, MSA medians are computed from tracts/block groups with centroids falling within the MSA. This approach continually incorporates new areas into the MSA, ensuring that newly developed regions, capable of exerting an upward influence at the MSA level, are not overlooked. As these variables are reported in bins, I assign midpoints except for the highest bin, which is capped. Subsequently, I calculate medians using these midpoints and bin-specific housing unit/family counts as weights.

[19]Family income is defined as family income in the previous year. It is only available since 1950.

neighborhoods falling below the MSA mean in terms of key socio-economic indicators and that most of them continued to lag behind the MSA means in 2015.

Overcoming the identification challenges caused by the HOLC grading process requires comparing similar areas that faced different levels of credit restrictions. Motivated by the HOLC grades, only C-graded areas constitute the control group since they were the areas considered more similar to D-graded ones and to be in the previous step before converging to a D zone. Evidence in Appendix Tables 8.3.8 and 8.3.9 provides additional supporting evidence showing the D and C neighborhoods have the smallest differences in population, housing, and income variables in 1940 and 1950. The D-C comparison has the additional advantage of capturing a discrete jump in the credit policy: from complete credit restriction (D) to conservative lending (C) as mentioned by Krimmel (2018).

The D-C comparison is performed at two different levels. The first level restricts the analysis to neighborhoods within the same MSA.[20] Adding MSA fixed effects to the diff-in-diff equations controls for MSA unobserved and time-invariant characteristics, while year fixed-effect control for common time-trends. The second level imposes further restrictions since, even within an MSA, areas may evolve differently subject to other unobserved factors. To overcome the concern I follow the usual approach in redlining literature that restricts the analysis to adjacent D-C neighborhoods sharing the longest borders, similar to Aaronson, Hartley, and Mazumder (2021) and Krimmel (2018).[21] D-C bordering areas represent sudden changes in grades but, being adjacent areas, they share the same unobservable and local characteristics, and thus, the grade assignment is *as good as* random.[22] The procedure resembles regression discontinuities usually employed in the education economics literature that exploit falling a different sides of a grade cut-off.[23] In the redlining setup, there is a similar grade cut-off between receiving the worst credit rating (D) and the second worst (C), and the threshold from the cut-off is defined geographically by being adjacent and sharing the longest border in a similar fashion as in border regression discontinuity. Adding border-pair fixed effects guarantees any common and time-invariant unobservable factor for the pair of neighborhoods is controlled for.[24]

One requirement for the empirical design to isolate the effect of redlining is that there is a sharp variation in the assigned grade at the border, but neighborhood characteristics change gradually. This requirement guided the data construction. As discussed in Section 3, data is

[20]The term *city* is used to reference the maps designation of cities. These surveyors' definitions of *cities* are cumbersome since they tend to divide areas in different maps (i.e., the 5 boroughs of New York). Hence, HOLC neighborhoods are matched to the corresponding MSA (2010 definition) to avoid these situations.

[21]The choice of neighboring areas on the basis of border length is made in this paper because, given the irregular shapes of HOLC neighborhoods, using centroids or coordinates as Krimmel (2018) does not allow to make a meaningful restriction.

[22]Appel and Nickerson (2016) follow a similar approach but comparing D areas to any other graded adjacent area. Their strategy exploits very different policies but also compares very different areas, threatening its validity

[23]See for instance Ost, Pan, and Webber (2018).

[24]Aaronson, Hartley, and Mazumder (2021) follow a more strict design and construct small buffers at each side of the C-D border. However, following such a procedure would not allow the incorporation of amenities in the analysis since it is unlikely to find geographic variation at small distances.

constructed at the original HOLC graded neighborhood level to preserve the original sharp variation in grades, but by assigning Census units to graded neighborhoods, unobservable neighborhood characteristics change gradually at the border.

The formal goal is to estimate the persistence of redlining using a diff-in-diff with two dimensions: redlining and the passing of the CRA. Let y_{imt} be the relevant dependent variable in HOLC neighborhood i at MSA m in year t, R_i be the redlining grade (1 if D-graded, 0 if C-graded) and $Post^{1977}$ represent the passing of the CRA (1 from 1980 onward, 0 until 1970), then the following equation estimates the average persistence of redlining:

$$y_{imt} = \beta_0 + \beta_1 R_i + \beta_2 (R_i \times Post^{1977}) + \alpha_{im} + \gamma_t + \epsilon_{imt} \tag{1}$$

where α_{im} represents either MSA fixed effect or border-pair fixed effects, γ_t are year fixed effects, and ϵ_{imt} is the error term.[25] In this regression, the coefficient of interest would be β_3: it reflects the catching up between D and C areas after the outlawing of redlining.

The estimates of Equation 1 at the within MSA and border-pair are shown in Table 1 and Table 2 respectively. Both tables lead to the same conclusions. Focusing on the first row, the coefficient for being D-graded is negative and statistically significant: $\beta_1 < 0$ in Equation 1. This shows that, during the years of redlining (1940-1970), there were negative significant gaps for D areas, compared to their C neighbors. The coefficient of the interaction between D-grade and the passing of the CRA is, however, positive and strongly significant: $\beta_2 > 0$ in Equation 1. This coefficient indicates that, after the removal of redlining, there is some degree of convergence for all the variables. However, adding up the two coefficients (i.e., $\beta_1 + \beta_3$ in Equation 1) shows that the D-C gaps are still present after the removal of redlining. Hence, the effects of redlining do not disappear and are persistent over time. Complementing the results of Section 3 on the magnitude of errors in estimation induced by assigning grades to tracts, Appendix Tables 8.3.15 and 8.3.16 provide additional evidence by estimating Equation 1 with a D-C sample in which each 1940 tract is assigned the HOLC grade with the largest spatial overlap. Comparing both sets of results implies that assigning grades to Census units biases downward average persistence, consistent with the experiment of Section 3.

From Table 1 and 2, average persistence can be computed by taking the ratio of the average gap after the passing of the CRA and the average gap during redlining (i.e., $\frac{\beta_1+\beta_3}{\beta_1}$ in Equation 1). Given that the estimates of the within MSA comparison can be biased in the presence of local unobserved factors, focusing on the border-pair results, shows that 53% of the gap in the white population, 32 % in home values and 72% of income persists after outlawing redlining.

[25] Note that, in my specifications with border-pair fixed effects, MSA fixed will be fully absorbed by the border-pair ones. Moreover, since year fixed effects are introduced, the variable $Post^{1977}$ would be colinear to these fixed effects. Given the data construction process, the number of observations per decade and MSA is relatively low, and hence MSA-year fixed effects to control for time-trends cannot be included since there is not enough variation to estimate them. For the same reason, standard errors cannot be clustered at the MSA-year level since clustering requires having enough observations per cluster. As a result, to take into account spatial correlation, the standard errors are clustered at the Census division-year level.

Table 1: Persistent effects of redlining, D-C neighborhoods in the same MSA

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-13.40***	-18.76***	-11.46***
	(1.12)	(1.33)	(0.74)
D-graded × Post^{1977}	3.81***	11.27***	3.14***
	(1.38)	(1.82)	(0.92)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.37	0.24	0.28
Adjusted within R^2	0.04	0.07	0.07
Average Persistence	72%	40%	73%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. The *Post*^{1977} period is 1980-2015. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 2: Persistent effects of redlining, bordering D-C neighborhoods

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-8.23***	-10.75***	-6.13***
	(0.64)	(0.91)	(0.44)
D-graded × Post^{1977}	3.88***	7.34***	1.71***
	(1.00)	(1.14)	(0.59)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	62.36	38.98	39.01
Observations	11,030	10,925	9,798
Adjusted R^2	0.73	0.53	0.63
Adjusted within R^2	0.03	0.04	0.05
Average Persistence	53%	32%	72%

Notes: All columns contain border-pair and year fixed effects, so coefficients are estimated on the basis of within D-C pairs. The *Post*^{1977} period is 1980-2015. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Figure 1: Timing of the effects

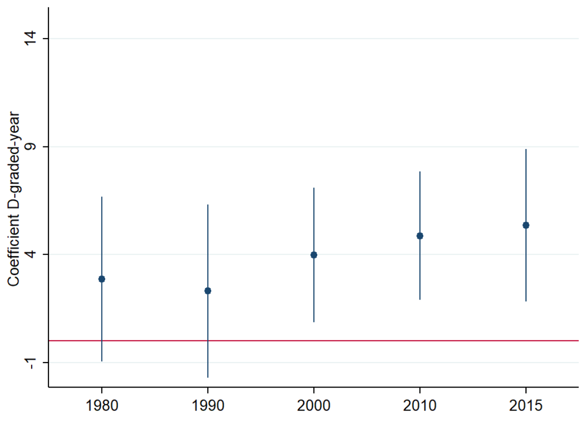


Figure A: White Share

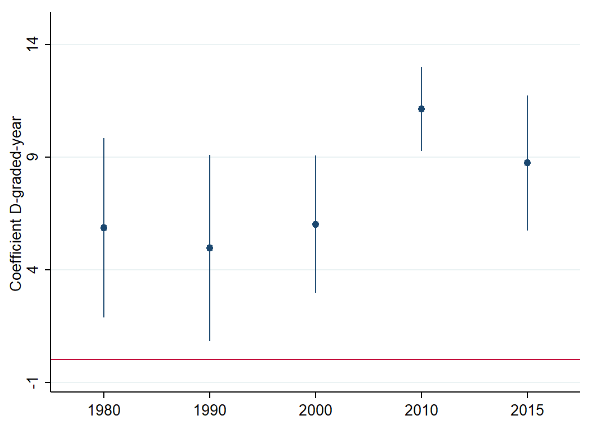


Figure B: Home Values

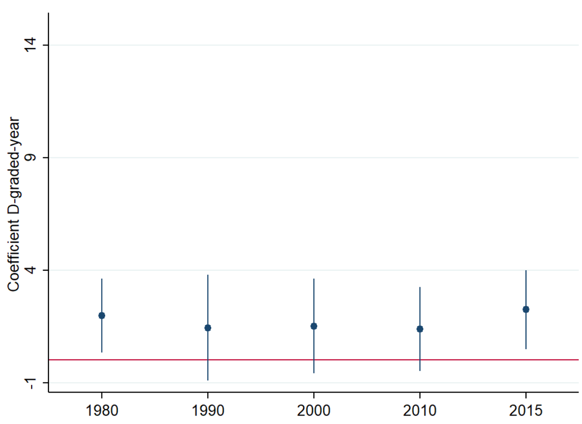


Figure C: Family Income

Notes: this figure shows the coefficient β_3 of estimating: $y_{imt} = \beta_0 + \beta_1 R_i + \beta_2 Post^{1977} + \beta_3(R_i \times Post^{1977}) + \alpha_{ij} + \epsilon_{imt}$, where $Post^{1977}$ takes values zero between 1940-1970 and a different value for each decade after the removal of redlining. The standard errors of the regression were clustered by Census division-year. Estimated on the bordering D-C sample. Dependent variable in Figure A is the share of white population; percentage of housing units on and above the MSA median home value in Figure B; percentage of families on and above the MSA median family income in Figure C.

To discern the timing pattern behind the effects, I modify Equation 1 by letting the $Post^{1977}$ dummy take value zero for the redlining year (1950-1970) and a different value for each decade afterwards[26]. The interaction between D-grade and the new post variable will show the yearly change in the gap, relative to the years when redlining was legal. The estimated coefficients of this interaction for each variable and the 95% confidence intervals are shown in Figure 1. Convergence in the share of white population occurs gradually and spreads over time. However, for home values and income the effects of removing redlining occur as soon as it is prohibited because the 1980 coefficient is positive and statistically significant.[27] For home values most of the convergence is happening in recent years, while for family income the trend is almost flat until 2015. [28]

5. The role of water and parks in the legacy of redlining

The findings in Section 4 are consistent with redlining literature and capture average persistence. However, estimating average effects makes it difficult to conclude that redlining still affects all neighborhoods to the same degree: the persistence may only occur in certain areas or under certain conditions.

This section departs from this conventional perspective by exploring if water and park amenities mitigate the redlining effects. Introducing water and park amenities in the framework is motivated by the literature showing they determine neighborhood long-run trajectories and explain the persistent spatial inequalities. The relationship is modeled by introducing an additional dimension representing natural amenities (A_i) in Equation 1:

$$\begin{aligned} y_{imt} = &\beta_0 + \beta_1 R_i + \beta_2 A_i + \beta_3 (R_i \times A_i) + \beta_4 (R_i \times Post^{1977}) + \\ &\beta_5 (A_i \times Post^{1977}) + \beta_6 (R_i \times Post^{1977} \times A_i) + \alpha_{im} + \gamma_t + \epsilon_{imt} \end{aligned} \tag{2}$$

where all variables are defined as in Equation 1. The main coefficient of interest is β_6 since it will capture if the catching-up is faster for D-graded areas with water and parks natural amenities.

Table 3 and Table 4 show the results of estimating Equation 2 at the MSA and bordering level, respectively. The second and third rows reveal the coefficient for the impact of water and park amenities during the years of redlining. When redlining was in place, being nearby amenities increased the D-C gaps. At the border-pair level, C-graded neighborhoods with

[26] Instead of including year fixed effects, this variation adds the newly defined dummy $Post^{1977}$, which is virtually the same as the fixed effects

[27] Additional supporting evidence for is in Appendix Tables 8.3.17 and 8.3.18, where Equation 1 is estimated restricting the $Post^{1977}$ to 1980. The coefficient of the interaction with $Post^{1977}$ is only significant for home values and family income.

[28] For income, there is no such a clear time trend. However, this is not necessarily wrong and would be consistent with neighborhood change and household sorting. The passing of the CRA could be leading to faster effects in terms of population and values and, as these effects take place, they will affect the income of the families that decide to move to a neighborhood.

Table 3: Natural amenities mitigate persistence, D-C neighborhoods in the same MSA

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-13.06***	-15.63***	-10.14***
	(1.01)	(1.30)	(0.78)
D-graded × Post^{1977}	3.90***	8.33***	2.33**
	(1.48)	(1.86)	(1.00)
Water or park amenities	2.32***	2.11**	1.18**
	(0.64)	(1.03)	(0.52)
Water or park amenities × Post^{1977}	3.54***	0.87	1.73*
	(1.00)	(1.64)	(0.95)
D-graded × Water or park amenities	-0.64	-5.06***	-2.15***
	(1.16)	(0.86)	(0.64)
D-graded × Water or park amenities × Post^{1977}	-0.32	4.59***	1.18
	(1.67)	(1.38)	(0.94)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.28
Adjusted within R^2	0.04	0.07	0.08
Average Persistence Water or Parks	74%	38%	71%
Average Persistence No Water nor Parks	70%	47%	77%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. The *Post*^{1977} period is 1980-2015. Water or park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 20% of the area. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without water or parks Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 4: Natural amenities mitigate persistence, bordering D-C neighborhoods

	(1)	(2)	(3)
Dependent variables	% white	% housing units above MSA median home value	% families above MSA median family income
D-graded	-7.92***	-9.23***	-5.51***
	(0.98)	(1.05)	(0.67)
D-graded × Post^{1977}	2.44	5.55***	1.15
	(1.48)	(1.35)	(0.94)
Water or park amenities	-0.89	-1.86*	-1.65**
	(0.93)	(1.06)	(0.65)
Water or park amenities × Post^{1977}	2.50*	2.50	2.27**
	(1.30)	(1.62)	(1.04)
D-graded × Water or park amenities	-0.58	-2.63***	-1.11
	(1.17)	(0.79)	(0.70)
D-graded × Water or park amenities × Post^{1977}	2.55*	3.13**	1.06
	(1.49)	(1.37)	(1.06)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	62.36	38.98	39.01
Observations	11,030	10,925	9,798
Adjusted R^2	0.73	0.53	0.63
Adjusted within R^2	0.04	0.05	0.06
Average Persistence Water or Parks	41%	27%	67%
Average Persistence No Water nor Parks	69%	40%	79%

Notes: All columns contain border-pair and year fixed effects, so coefficients are estimated on the basis of within D-C pairs. The *Post*^{1977} period is 1980-2015. Water or park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 20% of the area. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without water or parks. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

amenities would also experiment with lower home values and family income. Given that C-graded near amenities within the MSA exhibit the opposite situation, the border-pair effect could be driven by spillovers from their D-graded pairs. The negative signs do not necessarily challenge the hypothesis that amenities can mitigate the legacy of redlining, as their impact on neighborhoods can change over time rather than represent other instances of path dependence.

The triple interaction at the border pair shows that water and park amenities significantly mitigate the persistence of redlining in population and housing values. The lack of effect on income is explained by the fact that neighborhood change is a medium-long run phenomenon and household sorting depending on income takes place on a longer period, when the share of white population and home values have already changed. In fact, the timing pattern for income suggests in Figure 1, which is actually driven by D-graded areas by water or parks is consistent with this process of household sorting. Within MSA, convergence is only significantly stronger for home values. The lack of effect for white share relates to the redlined cities having high proportions of black population or experimenting with black inflows during the period.

From the point estimates of Equation 2, the degree of persistence for areas with and without amenities can be estimated taking the ratio of the average gap within subgroup after the CRA to the same ratio during redlining (i.e., $\frac{\beta_1+\beta_3+\beta_4+\beta_6}{\beta_1+\beta_3}$ for areas with amenities and $\frac{\beta_1+\beta_4}{\beta_1}$ for areas without). The results in Table 4 imply that the presence of natural amenities is not enough to achieve full convergence but it reduces persistence to 41% in population and 27 % in home values. This contrasts with the average persistence in Section 4 of 53% and 32% as observed in Section 4, and 69% and 40% for areas lacking such amenities.

Robustness

To assess the robustness of the results, I perform a series of tests shown in the Appendix. First, I estimate Equation 2 using placebo data. Placebo data is obtained by creating a random sequence of HOLC grades while keeping the grades' proportions in the entire sample. After defining the placebo grades, the same adjoining-longest border criterion of the paper determines the placebo D-C pairs. Results are shown in Appendix Table 8.3.19. None of the placebo D-grade coefficients are statistically significant, reinforcing the validity of the results. Significant effects are only found for natural amenities, which would have a positive and statistically significant effect after the passing of the CRA and a negative one during the redlining years. This finding supports the hypothesis that natural amenities impact neighborhood trajectories and that their effect can change over time.

The 20% threshold to define water and park amenities was chosen to balance capturing meaningful natural amenities without concerns of sample selection. Two strategies assess the robustness of this definition: using different thresholds or implementing a new definition capturing the same situations (i.e., meaningful but not restrictive).[29] To approximate the second situation, water and park amenities are redefined as the situation in which the share

[29] Results using a 10% and a 30% threshold can be found in Appendix Tables 8.3.22,8.3.23, 8.3.24 and 8.3.25.

of a neighborhood covered by any water feature or park is above the MSA median coverage for that feature, weighted by neighborhood area. The new definition captures meaningful amenities since they are above the median for the MSA and avoids selection issues since it considers all areas above the median. Appendix Tables 8.3.20 and 8.3.21 show the results with this definition. The main conclusions remain unchanged with the new definition except for the absence of a significant interaction between D-graded, natural amenities and the CRA in the share of the white population. Because the previous result was marginally significant and driven by recent years, changing the definition of water and park amenities may affect it.

6. Moulding neighborhood geography: waterfront renovations

The previous results indicate that water and park amenities mitigate the persistence of redlining, aligning with the literature documenting amenities as sources of persistent spatial differences. The following section departs conceptually from that literature by showing that natural amenities are not immutable but can be shaped through human intervention. While water and park amenities have a static and permanent component –i.e., their location – they have other aspects that can change. For instance, accessibility to water amenities and their utility improves by creating waterfront promenades, rehabilitating abandoned structures, or with brownfield cleanups.

The analysis in this section focuses on waterfront redevelopment plans that have occurred since the 1970s. These plans targeted former industrial or commercial zones left abandoned, polluted, and inaccessible due to shifts in industrial locations. Examples include Boston's North End, whose waterfront was once a major commercial and industrial area before being abandoned in the 1960s and 1970s. The Baltimore Inner Harbor followed a similar path, losing relevance after the introduction of container ships, as they could no longer dock there due to their size. City authorities established a series of redevelopment plans in these areas that included rehabilitating abandoned wharves and structures and creating and improving waterfront access (i.e., the creation of the Christopher Columbus Waterfront Park in North End). Both areas redeveloped quickly as a result of these strategies. These two stories illustrate that modifying water amenities is feasible and strongly impacts neighborhood trajectories. Following the success of Baltimore and Boston, other US cities adopted similar strategies to redevelop former industrial waterfronts.

In the same fashion as in the previous sections, the relationship between the persistence of redlining, the presence of water amenities, and their modifications can be expressed by adding an additional dimension to the diff-in-diff:

$$
\begin{aligned}
y_{imt} = & \beta_0 + \beta_1 R_i + \beta_2 A_i + \beta_3 (R_i \times A_i) + \beta_4 (A_i \times W_i) + \beta_5 (R_i \times A_i \times W_i) + \beta_6 (R_i \times Post^{1977}) + \\
& \beta_7 (A_i \times Post^{1977}) + \beta_8 (A_i \times W_i \times Post^{1977}) + \beta_9 (R_i \times A_i \times Post^{1977}) + \\
& \beta_{10} (R_i \times A_i \times W_i \times Post^{1977}) + \alpha_{im} + \sum_k \beta_k P_{imt} + \gamma_t + \epsilon_{imt}
\end{aligned} \tag{3}
$$

where W_i is an indicator for waterfront redevelopment projects, A_i captures water amenities and P_{imt} are the diff-in-diff counterpart for the presence of parks. The rest of the variables are defined as in the previous equations.[30] Because modifications only happen in areas with water and park amenities, only the interactions between A_i and W_i appear. The coefficient β_{10} represents the catch-up for areas with modified waterfronts compared to the convergence for areas with unmodified water amenities.

The results of estimating Equation 3 within MSA are shown in Table 5.[31] While the interaction between D-graded, unmodified amenities and the outlawing of redlining is no longer significant, except for home values, the coefficient with modifications is large and significant. The implication here is that the mitigation of persistence is not a universal outcome for all amenities; rather, the driving force behind these effects is the modified and revitalized amenities. In fact, for neighborhoods with waterfront modifications persistence gets reduced to -182% in white population, 12 % in home values, and -24% in family income, while for areas without improved waterfronts persistence is still 66% in population, 64% as in home values and 82% in family income. Notice that the negative persistence in home values and family income implies D-graded neighborhoods with modified waterfronts have largely overcome the legacy of redlining.

Robustness

Given that these waterfront modifications have been occurring since the 1970s, the results in Table 5 could capture the tendency of natural amenities to change over time rather than the effect of the modifications. To exclude the possibility, Equation 2 is estimated by adding natural amenities-year fixed effect to eliminate variation generated by these tendencies. Appendix Tables 8.3.26 and 8.3.27 show that the previous results remain unchanged even after including these fixed effects to absorb the time-trends.[32] Hence, adding an amenities-year trend does

[30]Modifications in this definition do not account for the timing. Since these modifications are relevant after the 70s, coefficients that do not interact with the $Post^{1977}$ variable will capture the situation of areas that will experiment with a waterfront redevelopment but have not been modified yet.

[31]Given that modified waterfronts were usually industrial or commercial areas that separated from the rest of the city, affected neighborhoods also tended to be separated or surrounded by D neighborhoods since these areas were the oldest part of the cities, inhabited by low-income population working on those industries and also because industrial and business *encroachment* was considered an adverse influence for surveyors and were associated with the worst grade. As a result, this equation can only be estimated within MSA.

[32]Notice that in this regression the interaction between natural amenities and $Post^{1977}$ is not included since it would be collinear to the amenities-year fixed effect.

Table 5: Waterfront modifications drive the effect of water amenities

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income income
D-graded	-12.53***	-14.75***	-9.95***
	(1.00)	(1.33)	(0.78)
D-graded × Post^{1977}	4.17***	8.19***	2.72***
	(1.43)	(1.75)	(1.00)
Water amenities	1.68*	1.23	1.99**
	(0.99)	(1.12)	(0.80)
Water amenities × Post^{1977}	5.38***	1.69	0.79
	(1.77)	(1.92)	(1.16)
D-graded × Water amenities	3.71**	-1.66	0.80
	(1.47)	(1.25)	(1.05)
D-graded × Water amenities × Post^{1977}	-1.20	-2.24	-1.07
	(2.83)	(2.10)	(1.57)
Water amenities × Modification	2.88	-0.20	-4.28*
	(2.54)	(3.15)	(2.52)
Water amenities × Modification × Post^{1977}	-5.12	4.15	0.53
	(5.26)	(4.86)	(4.61)
D-graded × Water amenities × Modification	4.08**	-4.64	-1.41
	(1.66)	(6.77)	(2.83)
D-graded × Water amenities × Modification × Post^{1977}	10.40***	12.64	11.46**
	(3.00)	(7.75)	(4.68)
Area FE	MSA	MSA	MSA
Park controls	YES	YES	YES
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.29
Adjusted within R^2	0.05	0.08	0.08
Average Persistence Modified	-182%	12%	-24%
Average Persistence Unmodified	66%	64%	82%
Average Persistence No Water nor Parks	67%	44%	73%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. $Post^{1977}$ is defined from 1980-2015. All columns control for parks (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the parks) and its interactions with being D-graded and $Post^{1977}$. Water amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features cover at least 20% of the area. Modification is an indicator for waterfront redevelopment projects (1 if the neighborhood falls within the 500 meter buffer around the project, 0 otherwise). Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

not absorb waterfront modifications.[33]

The definition of modifications used in Table 5 is static. By not considering the time when the modification happens, it pools modified and unmodified areas together. To explore the robustness of the results, I estimate a variation of Equation 3 using the same spatial overlap criterion but adding the timing so that they only appear as they happen. Given that these revitalization projects occurred only after the 70s,[34] the regression does not include modification variables not interacted with the $Post^{1977}$. The results with the new definition in Table 8.3.29 are nearly identical to the ones with the static definition.

A natural concern behind the results is that areas with modifications may systematically differ from the rest just before the redevelopments. To provide evidence on the pre-trends, I adopt an event-study design setting as reference date the previous decennial Census period to the modification . The graphs accompanying the event-study estimation for home values, which is the variable that can change faster to modifications, are shown in Figure 6. 10 years before the modifications, modified neighborhoods were not doing systematically worse than the rest of the neighborhoods within the MSA.

7. Greening the legacy of redlining.

Previous findings have revealed the existence of potential interventions that can alter water amenities and reverse the effects of redlining. These waterfront redevelopments, however, are viable primarily in neighborhoods located proximate to water features. Given the geographic limitations of implementing such policies, the potential for mitigating persistence remains restricted for areas distant from water bodies. In light of this, the following section delves into the role of an amenity with broader geographical applicability—tree canopy—in mediating the effects of redlining.

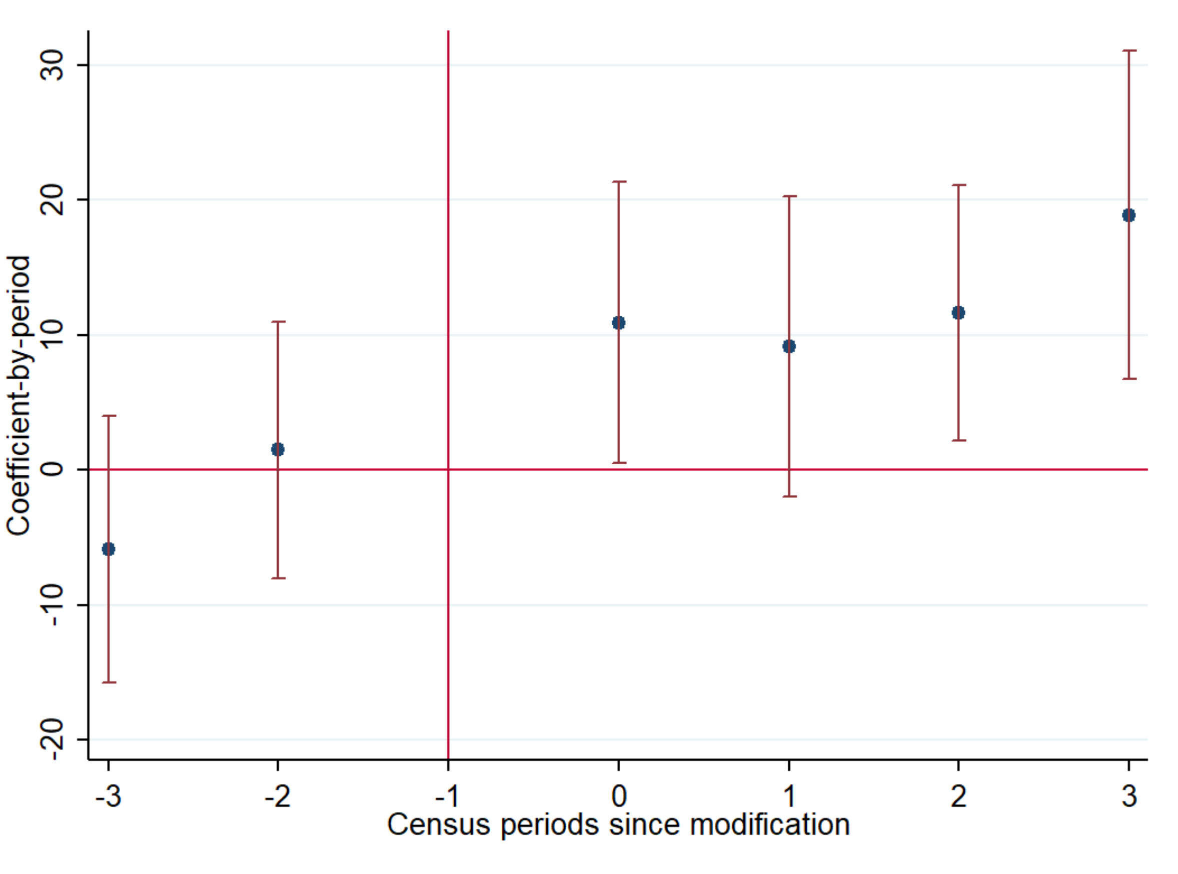
Due to data limitations, exploring the effects of tree canopy changes on neighborhood outcomes requires a different strategy. Modifying Equation 1 to incorporate changes in tree coverage in the same fashion as for water and park amenities and waterfront modifications is unfeasible given the lack of high-resolution aerial imagery (i.e., $1m^2$ pixels or less) before 2003. However, it is feasible to take a snapshot and estimate how outcomes in 2015 differ for D-graded neighborhoods depending on the experimented changes in the tree canopy. Formally, the goal is to estimate:

$$y_{im}^{2015} = \beta_0 + \beta_1 R_i + \beta_2 \Delta TC_i^{2015} + \beta_3 (R_i \times \Delta TC_i^{2015}) + \alpha_{im} + \epsilon_{im} \tag{4}$$

[33] Results in Table 5 are also robust to these fixed effects (not shown).

[34] The only exception would be Chicago Front Trail, to which I assigned 1964 because it was the only date found.

Figure 2: Home values and timing of waterfront modifications



Notes: the regressions in the figures controls for MSA and year fixed effects, so coefficients are estimated on the basis of all neighborhoods within MSA. All regressions control for HOLC grades and a dummy taking value one for areas where at least 20% of the area is covered by 500 meter buffers around water or parks. The standard errors of the regression were clustered by Census division-year. Dependent variable is the percentage of housing units on and above the MSA median home value. The modifications dates are rounded to the nearest Census period.

where all variables are defined as before and ΔTC_i^{2015} represents the growth rate of detected tree pixels between the 2000s-2015 (i.e., $\frac{TC_i^{2015}-TC_i^{2000s}}{TC_i^{2000s}}$). [35] The coefficient of interest in Equation 4 is β_3. It captures how the D-C gap changes for a 100% increase in tree coverage.

However, estimating Equation 4 with OLS would result in biased estimates because changes in urban trees (i.e., ΔTC_i^{2015}) are potentially endogenous. Equation 4 may suffer from reverse causality, and increases in tree canopy could be either a cause or a consequence of better neighborhood outcomes. There is also potential for omitted variable bias, and unobserved events and local interventions associated with changes in urban green spaces (i.e., new constructions or sidewalks and street renewals) could be biasing the estimates. Moreover, even though the tree detection algorithm achieves an average precision of 90%, measurement error is unavoidable, and some pixels will be incorrectly labeled. Furthermore, there is evidence that differences in tree canopy are related to redlining, with D-graded areas having lower levels of vegetation cover (Locke, Hall, Grove *et al.*, 2021; Nardone, Rudolph, Morello-Frosch, and Casey, 2021; Namin, Xu, Zhou, and Beyer, 2020). Similarly, redlining influences the evolution of tree coverage because D-graded areas tend to meet the criteria for priority plantation sites (i.e., high imperviousness, urban heat islands, social and economic inequality).[36] Appendix Table 8.3.12 provides additional evidence showing that (1) D neighborhoods are the ones with the lowest share of tree pixels, (2) together with C areas, they are the ones below total average canopy cover and (3) despite of a general increase in tree cover for all neighborhoods, the highest increase takes place in redlined areas.

Nonetheless, the sign of the bias generated is unclear. An upward bias would be caused if it is gentrifying areas that demand more trees. In such a situation, the OLS coefficients would be larger than the instrumented ones. On the other hand, if trees are being disproportionately planted in areas lagging as an intervention to improve their situation, the OLS coefficients would be downward biased. Regardless of the sign of the bias, any of the situations imply that the OLS coefficients cannot be causally interpreted.

To address these endogeneity concerns, this paper employs a two-stage least squares approach and predicts changes in tree coverage with changes in exposure to exotic tree plagues.[37] The plagues behind the instrument were the deadliest in 2015 according to Fei, Morin, Oswalt, and Liebhold (2019). To further ensure relevance and exogeneity, two plagues

[35] Due to computational constraints, changes in detected trees can only be measured between two time periods. The first period is labeled the 2000s since the first available data year depends on the states and ranges between 2003-2007. For a detailed explanation of data construction, see Section 3.

[36] See the evidence provided by Hoffman, Shandas, and Pendleton (2020) on redlining, imperviousness, and urban heat islands.

[37] Compared to native plagues, exotic ones represent a greater threat due to (i) the limited co-evolution between hosts and plagues that reduces host resistance (Tubby and Webber, 2010) and (ii) the lack of native enemies that facilitates the spread upon arrival (Aukema, Leung, Kovacs, Chivers, Britton, Englin, Frankel, Haight, Holmes, Liebhold *et al.*, 2011). Some examples include the Gypsy Moth, accidentally released in the 1860s, and that between 1920-2002 defoliated over 95 million acres (Coleman, Haavik, Foelker, and Liebhold, 2020). The arrival of the Dutch Elm Disease (DED) to Ohio in the 1930s caused similar consequences killing 56% of the original northeastern elms in the next 40 years. Other examples include the Hemlock Wolly Adelgid, the Asian Longhorned Beetle, and the recent Emerald Ash Borer. On average, host mortality occurs within 4-10 years of infection of these plagues.

with different mortality and management strategies are excluded.[38] The first one is the Gypsy Moth since host mortality occurs only after successive defoliation, which is unobservable using available data, and recent management strategies have focused on mating disruption to slow its spread. The other excluded plague is the White Pine Blister Rust, a pathogen whose relatively long time of latent infection, along with the fact that it spreads through infected ribes and not from tree-to-tree implied tree removals were ineffective ways of managing the disease (Maloy, 2003). For the considered plagues, chemical treatments, when available, are usually preventive, typically need to be repeated annually or biannually, and due to their costs and ecological impact, are generally only recommended for high-value ornamental trees. As a result, management strategies typically include a combination of host removals and preventive treatments when these costs are not expected to rise in an eventual infection.[39]

Non-native tree plagues are exogenous shocks whose management requires the removal of affected trees and their replacement (Aukema *et al.*, 2011; Hudgins, Koch, Ambrose, and Leung, 2022). Because the size, species, condition, and location of removed trees determine replacement in an attempt to maintain their value, replacements may not be on a one-to-one basis.[40] Typically, medium-to-large basal area trees, which are the most vulnerable to removal, must be substituted by more than one smaller tree. Hence, variation in plagues stemming from lengthier exposure and new plagues' arrival will trigger an exogenous change in the canopy that can be used as an instrumental variable.

Using the data on pests and potential hosts distribution described in Section 3 involves assuming all neighborhoods within a given county infested by plague j will also be infested if they contain hosts for pest j and that exposure increases with the area of potential hosts. Letting PH_{ij} be the share of plague j potential hosts basal area to detected trees in 2000 in neighborhood i and Y_{ij}^{t} the years since the detection of plague j in that area, changes in plague exposure in neighborhood i between 2000-2015 are defined as follows:

$$\Delta PlagueExposure_i^{2015} = \sum_{j=0}^{j=5} PH_{ji}^{2000} \times \Delta Y_{ij}^{2015} \tag{5}$$

Equation 5 captures different sources of variation in exposure: variation in the share of

[38]Additionally, four plagues are not detected in the data used. These plagues are the Green Spruce Aphid, the Laurel Wilt, the Sudden-Oak Death, and the Port-Orford-cedar root disease.

[39]Costs increase with the tree basal area (area covered by the stems) whose evolution is closely related to the age of the tree. Typically, young trees and old trees tend to grow slower than middle-aged. Therefore, medium-age trees with medium-to-large basal areas are removed, while young and old trees may be preventively treated.

[40]An example can be seen in the New York City Department of Parks & Recreation regulations: `https://www.nycgovparks.org/rules/section-5`. Other practical examples are available at the Tree Plantation guidelines of Arlington: `https://www.arlingtonva.us/Government/Programs/Building/Resources/Tree-Replacement`. Research suggests that tree replacement based on leaf area would range from 13.7 per large removed tree to 3.3 per small removed tree (Nowak and Aevermann, 2019). Moreover, new plantations can employ non-host species or genetically resistant hosts (i.e., the Pacific hemlock is immune to plagues affecting the Atlantic variant). In fact, for endemic pests, current research is trying to develop host-resistant species rather than treatments.

trees potentially infected (PH_{ji}^{2000}), [41] and variation in the years of exposure (Y_{ij}^{2015}). Since particular species may be endogenously allocated to neighborhoods, considering all j plagues combined strengthens the exogeneity of the instrument by capturing susceptibility to any plague rather than to a specific one. Including exposure times captures that the effect on tree replacements rises with longer exposures as more hosts are affected. Similarly, for replaced trees, there must be some time lag before their size is large enough to be detected in imagery. Appendix Figures 8.3.1 and 8.3.2 display the county distribution of plagues and an illustration of the share of potential hosts of one of the plagues, the Emerald Ash Borer, in Chicago's neighborhoods respectively

Then, the first stage equation is defined as:

$$\Delta TC_i^{2015} = \alpha_0 + \alpha_1 R_i + \alpha_2 \Delta PlagueExposure_i^{2015} + \alpha_3(R_i \times \Delta PlagueExposure_i^{2015}) + \alpha_{im} + u_i \tag{6}$$

where all variables are defined as in the text. Notice that there was only one endogenous variable, ΔTC_i^{2015}, in the OLS equation. R_i is included since it will appear in the second stage equation, and if ΔTC_i is, in fact, related to R_i but not controlled for, the error term of Equation 6 will be correlated with R_i and will bias the estimations. Since R_i is uncorrelated with u_i, so is the interaction between R_i and $\Delta PlagueExposure_i^{2015}$. Adding the interaction has the additional advantage of controlling for possible concerns regarding heterogeneity in plague effects and management. [42] Decomposing Equation 6 into fitted values ($\widehat{\Delta TC_i}$) and an error term (ν_i) and plugging this decomposition in Equation 4 yields:

$$y_{im}^{2015} = \beta_0 + \beta_1 R_i + \beta_2 \widehat{\Delta TC_i^{2015}} + \beta_3(R_i \times \widehat{\Delta TC_i^{2015}}) + \alpha_{im} + \zeta_i \tag{7}$$

where $\zeta_i = \beta_2 \nu_i + \beta_3(R_i \times \nu_i) + \epsilon_{im}$. Estimating this equation would be problematic if any of the regressors is correlated with the error ζ_i. However, notice that $\widehat{\Delta TC_i^{2015}}$ would be, by construction, orthogonal to both ν_i, ϵ_{im} and R_i, and hence uncorrelated with the error. Similarly, R_i will also be uncorrelated to u_i since it is included as a regressor in the first stage and is thus orthogonal to ϵ_{im}. The only potential concern would be the correlation between R_i and the term $R_i \times \nu_i$, but since R_i is orthogonal to ν_i and R_i^2 is R_i (i.e., it is a dummy variable), there is no correlation between regressors and ζ_i and hence Equation 7 can be estimated safely.[43]

The underlying assumption behind the use of exotic pests as instruments is that they have an equal effect on all neighborhoods. This assumption is not unrealistic since infected, dead, or at-risk trees, private or public, are equally likely to be removed regardless of their location as long as they share similar ages and basal areas. Since this paper focuses on similarly old

[41] The normalization with pixels detected in the 2000s is done because the increases in tree canopy are also defined relative to these pixels.

[42] For instance, given the evidence in Hoffman, Shandas, and Pendleton (2020) on redlining areas suffering from urban heat islands, it could be that trees in redlined areas are subject to more stress and therefore be more likely to die from plagues. These effects, if they exist, will be accounted for by the interaction of both variables.

[43] $Cov(R_i, R_i\nu_i) = E(R_i^2\nu_i) = E(R_i\nu_i) = E(R_i)E(\nu_i) = 0$.

neighborhoods, D and C-graded ones, the overall distribution of tree canopy should be akin since they have experienced the same shocks. As a result, changes in exposure to plagues should have comparable effects in both neighborhoods. Other potential threats could be that mortality rates are endogenous due to the environmental stresses for trees caused by redlining or that replacements are endogenous and D-graded areas receive lower or slower-growth replacement trees. Importantly, these hypotheses can be checked empirically by looking at the estimates of α_1 and α_3 in Equation 6.

Given the spatial correlation between species distribution and plagues for neighboring areas, estimating Equation 6 and 7 at the border-pair level is unfeasible since there would not be enough variation in plague exposure to predict tree canopy after adding border-pair fixed effects.

Table 6 displays, on Panel A, the results of estimating the first stage equation (Equation 6) and, on Panel B, the results of the second stage (Equation 7). The first stage results show that increases in plague exposure lead to significant increases in the tree canopy. To simplify the interpretation of Δ Plague Exposure$_i^{2015}$, a standard deviation increase in plague exposure leads to 237 pp higher increases in tree coverage. Column (2) of Panel A also controls for natural amenities, their modifications, and the interactions with redlining since these features could correlate with the observed changes in the tree canopy. The results remain unchanged even after adding these additional controls. Moreover, results in column (1) do not show significant heterogeneity in the effect of plague exposure for D and C neighborhoods, reinforcing the exogeneity of the instrument.

Comparing the second stage results in Panel B with the OLS results in Table 7 shows that the OLS estimates are downward biased. Moreover, while there are no significant effects for the interaction between D-graded and changes in tree coverage with OLS, the interaction becomes significant and positive for the white population and family income using the two-stage least squares strategy. Doubling tree canopy reduces the demographic and income gaps by 40% (i.e., $1 - (\beta_1 + \beta_3)/\beta_1$ in Equation 7). The lack of a significant effect on home values at the neighborhood level is consistent with the literature on the hedonic analysis of trees. Because the impact of trees on property prices decays with distance, observing only medians of values at the neighborhood level can offset the effect.

The difference between IV and OLS estimates is consistent with Baum-Snow (2007) and Duranton and Turner (2012), suggesting that the increases in tree coverage result from policy interventions in areas that have not converged to achieve the catching-up. The existence of tree plantation and regreening initiatives in low-income areas further sustains this hypothesis. For instance, Groundwork USA, a network of approximately 20 local trusts, was founded in 1998 from a partnership between the National Park Service and the Environmental Protection Agency and is devoted to improving the environmental conditions of low-resource communities and reverting the legacy of poverty and discrimination through multiple greening initiatives. Similarly, the Environmental Tree Service in Portland has provided free street trees to low-income and under-served communities since 2008. Moreover, with the publication of the

Table 6: Greening redlining

Panel A: First stage

	(1)	(2)
Dependent variables	Δ Tree canopy	Δ Tree canopy
D-graded	0.844	0.635*
	(0.6590)	(0.3317)
Δ Plague Exposure	213.760***	214.204***
	(50.0954)	(50.3676)
D-graded × Δ Plague Exposure	-155.699	-149.762
	(155.3191)	(151.8108)
Δ 1 SD Δ Plague Exposure	2.37	2.38
F-stat $^{(\text{instrument})}$	18	18
F-stat $^{(\text{instrument \& interaction})}$	9	9
Area FE	MSA	MSA
Amenities and modifications		YES
Mean Dep. Var.	2.02	2.02
Observations	1452.00	1452.00
Adjusted R^2	0.15	0.15
Adjusted within R^2	0.07	0.07

Panel B: Second stage

	(1)	(2)	(3)
Dependent variables	% white	% housing units above MSA median home value	% families above MSA median family income
D-graded	-84.321***	-22.808	-45.730**
	(29.3970)	(18.9841)	(18.9578)
$\widehat{\Delta TC}$	0.456	0.082	0.006
	(0.5356)	(0.2444)	(0.1518)
D-graded × $\widehat{\Delta TC}$	33.976***	8.672	17.357**
	(12.8753)	(8.3026)	(8.2927)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	43.43	38.42	35.68
Observations	1,450	1,450	1,450

Notes: Panel A shows the results of the first stage equation, which regresses the experimented tree canopy increase on a dummy for being D-graded, the change in plague exposure and the interaction between both. Both columns include MSA fixed effects. Column (2) controls also for the presence of amenities modifications and the interactions with D-graded. The increase in tree canopy is computed as the growth of tree pixels between the two periods. Panel B shows the results from regressing the dependent variables in 2015 on the entire D-C sample on a dummy for being D-graded, the fitted values of the regression from Panel A and the interaction. All columns include MSA fixed effects. MSA without plagues are excluded. Standard errors are robust and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 7: OLS results

Dependent variables	(1) White share	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-5.090***	-2.439*	-6.428***
	(1.7973)	(1.3599)	(1.0497)
Δ Tree Canopy	0.845*	0.612**	-0.012
	(0.4436)	(0.3052)	(0.1329)
D-graded $\times$ Δ Tree Canopy	-0.680	-0.353	0.195
	(0.4386)	(0.3086)	(0.1477)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	43.43	35.68	35.68
Observations	1,448	1,448	1,448
Adjusted R^2	0.07	0.39	0.18
Adjusted within R^2	0.02	0.02	0.03

Notes: this table shows the results from regressing the dependent variables in 2015 on D-graded, the experimented increase in tree canopy and their interaction on the MSA D-C sample. Changes in tree canopy are defined as the increase in pixels detected as trees in between 2015 and 2000s. All specifications include MSA fixed-effects. MSA without plagues are excluded. Standard errors are robust and ***, **, * indicate significance at the 1, 5, and 10 percent.

HOLC maps, initiatives also started to focus explicitly on formerly redlined neighborhoods. For instance, the Southside ReLeaf association has been committed to reverting the environmental legacy of redlining in South Richmond since 2019.

Robustness

Given the high difference between OLS and IV estimates, I also estimate the reduced form of Equation 4 introducing changes in plague exposure directly. Results shown in Table 8 corroborate the previous finding: D-graded areas that experiment with higher exposure to plagues have higher shares of white population and family income. As in Table 6 there are no significant effects on housing values. Appendix Table 8.3.31 shows that these results remain unchanged even after controlling for natural amenities, modifications, and their interaction with redlining. Moreover, estimating the second stage controlling for natural amenities, modifications, and the interaction with redlining and using the fitted first-stage values of Column (2) in Panel A Table 6 does not lead to significant differences in the estimates, as shown in Appendix Table 8.3.32.

Table 8: Reduced form results

Dependent variables	(1) White share	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-7.667***	-3.272**	-6.763***
	(1.6963)	(1.3235)	(1.0856)
Δ Plague Exposure	97.479	17.510	1.265
	(114.4816)	(52.2514)	(32.4534)
D-graded $\times$ Δ Plague Exposure	1901.689**	490.767	1006.811**
	(754.6791)	(484.2337)	(482.3129)
Δ 1 SD Δ Plague Exposure	21.11	5.45	11.18
Mean Dep. Var.	43.43	38.42	35.68
Observations	1,450	1,450	1,450
Adjusted R^2	0.06	0.39	0.18
Adjusted within R^2	0.02	0.00	0.03

Notes: This table shows the results from regressing the dependent variables in 2015 for the entire D-C sample on a dummy for being D-graded, the experimented change in plague exposure and their interaction. All columns include MSA fixed effects. Changes in tree canopy are defined as the increase in tree detected pixels during the two periods with aerial imagery. MSA without plagues are excluded. Standard errors are robust and ***, **, * indicate significance at the 1, 5, and 10 percent.

8. Concluding comments

In response to the mortgage and housing crisis that followed the Great Depression, the New Deal administration undertook a series of reforms that had long-lasting consequences on these sectors. Among them are the creation of the Home Owner's Loan Corporation (HOLC) and the implementation of the *City Survey Program*. Under this plan, the HOLC graded neighborhoods in US cities with a population greater than 40,000 inhabitants to assess the risk of insuring mortgages supposed for the Federal Government in each neighborhood. The purpose was to design a system that would allow to guarantee and control of the value of the housing assets held by the government through mortgage insurance and refinancing. Neighborhoods could be assigned different grades (A-B-C-D) depending on the area transportation access, proximity to amenities, housing, and economic, demographic, and racial characteristics. The grades were then represented with different colors (green-blue-yellow-red, respectively) in the *Residential Security Maps*, commonly known as redlining maps. The appraisal criteria reflected the institutionalized racism of the period, leading to minority, black, and poor neighborhoods receiving the worst grade (D-graded). D-graded neighborhoods were deemed risky and hence deprived of federal housing credit until 1977, with the passing of the Community Reinvestment Act. The practice of systematically denying credit based on neighborhood characteristics is commonly known as redlining. The recent digitization of these maps (Nelson, Winling, Marciano, Connolly *et al.*, 2017)opened the path for the study on the long-term consequences of these discriminatory lending practices that show redlining has had persistent effects (Appel and Nickerson, 2016; Krimmel, 2018; Aaronson, Hartley, and Mazumder, 2021).

Utilizing the digitized redlining maps, Census data, and the location of water and park amenities for the 1940-2015 period, this paper identifies the relationship between the persistence of spatial inequalities and proximity to water and parks, leveraging the inequalities generated by redlining. Introducing waterfront revitalization projects departs from the traditional literature that considers geography as static by showing that water amenities also have a component that changes and can be molded through human intervention. In this way, this paper not only contributes to the above-quoted literature on the persistence of redlining but also to the literature that focuses on natural amenities as determinants of neighborhood outcomes (Rappaport and Sachs, 2003; Rappaport, 2007; Villarreal, 2014; Lee and Lin, 2018; Heblich, Trew, and Zylberberg, 2021). Given the limited geographic span of waterfront beautification projects, this paper also explores how changes in urban green coverage can mediate the effects of redlining. This paper also contributes to the literature on the effects of urban tree coverage with the development of new neighborhood panel data on tree canopy and a new instrumentation strategy to tackle the endogeneity behind changes in the tree canopy.

The empirical strategy overcomes the non-random grading by implementing a diff-in-diff comparing neighborhoods with the most severe credit restriction (D-graded, redlined) to nearby areas with the second worst grade (C-graded, yellowlined). Focusing on nearby D-C pairs leverages the new procedure to match redlining maps with Census data here developed, the Census-to-Redlining Constant Crosswalks. By assigning Census units to graded neighborhoods, the crosswalks preserve the original sharp variation in grade assignment while allowing unobservables to change gradually at the border. Results show that significant gaps between D-C areas have persisted decades after outlawing redlining, but the persistence is heterogeneous and diminishes in D-graded neighborhoods by water and parks. It is, however, only the water amenities that have been made accessible and improved to neighborhoods through waterfront revitalizations that drive the effects. Finally, the results also show that exogenous increases in tree coverage can completely close the D-C gaps in population and income.

But what is beyond the faster convergence due to waterfront renovations and increases in tree coverage? Contrasting stories with different implications could drive the results. On the one hand, if the original redlined communities are owners, they could capitalize on the increases in higher home values. In this way, residents' welfare would improve, not only because of the generated water and green amenities but also because higher home values would lead to higher tax revenue for local services. Therefore, waterfront beautification and re-greening strategies could effectively reverse the consequences of redlining for neighborhoods and individuals. On the other hand, it could be that the original residents who remained renters get displaced. Neighborhoods with improved waterfronts or greened areas may be experimenting with large influxes of white people. Increased demand for these neighborhoods would raise housing prices and rents. Eventually, the process would displace the original residents to areas with similar pre-convergence characteristics, or even worse if displacement is towards areas without water and park amenities, whether improved or not. As a result, despite successfully allowing D-graded neighborhoods to converge, the originally redlined communities would be worse off.

To approximate this question, the last four tables of the Appendix 8.3 re-estimate the equations for waterfront modifications and tree canopy but with dependent variables intersecting race, ownership, education, and income. Only D-graded neighborhoods with modified water amenities have experienced a significant decrease in the percentage of black owners and renters and an increase in the percentage of white owners and renters. Improved D-graded neighborhoods also have a higher proportion of white population with some college attendance. Interpreting the results of the Section on waterfront improvements in light of these two new findings suggests that the lower persistence was, in part, reflecting the gentrification of these neighborhoods.

Re-greening neighborhoods provoke similar results, decreasing the share of black owners and renters and increasing the white shares. In quantitative terms, a 100% exogenous growth proves sufficient to reduce by 40% the ownership and renting differences. However, the increases in tree coverage also result in higher high-income Black residents while they have no effects on the college education of residents.

This paper demonstrated that not all D-graded neighborhoods have remained degraded. Areas near water and park amenities can converge much faster. The heterogeneity documented here implies that redlining is not affecting all previously D-graded areas equally. As a result, policy interventions should focus on the ones that have remained degraded. Although this paper shows that waterfront and re-greening initiatives are very effective in revitalizing neighborhoods, it also highlights that these interventions are not necessarily associated with improvements for the original redlined communities.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder. 2021. The Effects of the 1930s HOLC "Redlining" Maps. *American Economic Journal: Economic Policy* 13(4): 355–92.

Almagro, Milena and Aradhya Sood. 2023. De jure versus de facto discrimination: Evidence from racial covenants .

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

Appel, Ian and Jordan Nickerson. 2016. Pockets of Poverty: The Long-Term Effects of Redlining .

Aukema, Juliann E., Brian Leung, Kent Kovacs, Corey Chivers, Kerry O. Britton, Jeffrey Englin, Susan J. Frankel, Robert G. Haight, Thomas P. Holmes, Andrew M. Liebhold *et al.* 2011. Economic impacts of non-native forest insects in the continental United States. *PLoS one* 6(9): e24587.

Baum-Snow, Nathaniel. 2007. Did Highways Cause Suburbanization? *The Quarterly Journal of Economics* 122(2): 775–805.

Bleakley, Hoyt and Jeffrey Lin. 2012. Portage and Path Dependence. *The Quarterly Journal of Economics* 127(2): 587–644.

Bosch, Martí. 2020. Detectree: Tree detection from aerial imagery in Python. *Journal of Open Source Software* 5(50): 2172.

Boustan, Leah Platt, Devin Michelle Bunten, and Owen Hearey. 2018. Urbanization in the United States, 1800-2000. In Louis P. Cain, Price V. Fishback, and Paul W. Rhode (eds.) *Oxford Handbook of American Economic History*. Oxford University Press".

Brueckner, Jan K., Jacques-Francois Thisse, and Yves Zenou. 1999. Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European economic review* 43(1): 91–107.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review* 106(4): 855–902.

Chyn, Eric and Lawrence F. Katz. 2021. Neighborhoods Matter: Assessing the Evidence for Place Effects. *Journal of Economic Perspectives* 35(4): 197–222.

Coleman, Tom W., Laurel J. Haavik, Chris Foelker, and Andrew M. Liebhold. 2020. Gypsy moth. *Forest Insect and Disease Leaflet 162.* US *Department of Agriculture, Forest Service. 20 p.* .

Crompton, John L. and Sarah Nicholls. 2019. The Impact of Park Views on Property Values. *Leisure Sciences* 0(0): 1–13.

Crossney, Kristen B. and David W. Bartelt. 2005. Residential security, risk, and race: The Home Owners' Loan Corporation and mortgage access in two cities. *Urban Geography* 26(8): 707–736.

Duranton, Gilles and Matthew A. Turner. 2012. Urban Growth and Transportation. *The Review of Economic Studies* 79(4): 1407–1440.

Faber, Jacob W. 2020. We Built This: Consequences of New Deal Era Intervention in America's Racial Geography. *American Sociological Review* 85(5): 739–775.

Fei, Songlin, Randall S. Morin, Christopher M. Oswalt, and Andrew M. Liebhold. 2019. Biomass losses resulting from insect and disease invasions in US forests. *Proceedings of the National Academy of Sciences* 116(35): 17371–17376.

FHA. 1936. Underwriting Manual: Underwriting Analysis under Title II, Section 203 of the National Housing Act.

Fishback, Price. 2014. Panel Discussion on Saving the Neighborhood: Part III .

Fishback, Price, Jonathan Rose, Ken Snowden, and Thomas Storrs. 2022. New Evidence on Redlining by Federal Housing Programs in the 1930s. *Journal of Urban Economics* : 103462.

Fishback, Price V, Jessica LaVoice, Allison Shertzer, and Randall Walsh. Forthcoming. The HOLC Maps: How Race and Poverty Influenced Real Estate Professionals' Evaluation of Lending Risk in the 1930s. Technical report.

Franco, Sofia F. and Jacob L. Macdonald. 2018. Measurement and valuation of urban greenness: Remote sensing and hedonic applications to Lisbon, Portugal. *Regional Science and Urban Economics* 72: 156–180. New Advances in Spatial Econometrics: Interactions Matter.

Greer, James. 2013. The Home Owners' Loan Corporation and the development of the residential security maps. *Journal of Urban History* 39(2): 275–296.

Han, Lu, Stephan Heblich, Christopher Timmins, and Yanos Zylberberg. 2021. Cool Cities: The Value of Green Infrastructure .

Hanlon, W.Walker and Stephan Heblich. 2022. History and urban economics. *Regional Science and Urban Economics* 94: 103751. Urban Economics and History.

Harriss, C. Lowell. 1951. History and Policies of the Home Owners' Loan Corporation. *NBER Books* .

Heblich, Stephan, Alex Trew, and Yanos Zylberberg. 2021. East-Side Story: Historical Pollution and Persistent Neighborhood Sorting. *Journal of Political Economy* 129(5): 1508–1552.

Hillier, Amy E. 2003. Redlining and the Home Owners' Loan Corporation. *Journal of Urban History* 29(4): 394–420.

Hillier, Amy E. 2005. Residential security maps and neighborhood appraisals: The Home Owners' Loan Corporation and the case of Philadelphia. *Social Science History* 29(2): 207–233.

Hoffman, Jeremy S., Vivek Shandas, and Nicholas Pendleton. 2020. The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas. *Climate* 8(1).

Holtan, Meghan T., Susan L. Dieterlen, and William C. Sullivan. 2015. Social Life Under Cover: Tree Canopy and Social Capital in Baltimore, Maryland. *Environment and Behavior* 47(5): 502–525.

Hudgins, Emma J., Frank H. Koch, Mark J. Ambrose, and Brian Leung. 2022. Hotspots of pest-induced US urban tree death, 2020-2050. *bioRxiv* .

Hynsjö, Disa M and Luca Perdoni. 2023. The Effects of Federal "Redlining" Maps: A Novel Estimation Strategy .

Jackson, Kenneth T. 1980. Race, Ethnicity, and Real Estate Appraisal: The Home Owners Loan Corporation and the Federal Housing Administration. *Journal of Urban History* 6(4): 419–452.

Jackson, Kenneth T. 1987. *Crabgrass frontier: The suburbanization of the United States*. Oxford University Press.

Kondo, Michelle C., SeungHoon Han, Geoffrey H. Donovan, and John M. MacDonald. 2017. The association between urban trees and crime: Evidence from the spread of the emerald ash borer in Cincinnati. *Landscape and Urban Planning* 157: 193–199.

Krimmel, Jacob. 2018. Persistence of Prejudice: Estimating the Long Term Effects of Redlining .

LaVoice, Jessica. 2019. The long-run implications of slum clearance: A neighborhood analysis.

Lee, Sanghoon and Jeffrey Lin. 2018. Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income. *The Review of Economic Studies* 85(1): 663–694.

Locke, D. H., B. Hall, J.M. Grove *et al.* 2021. Residential housing segregation and urban tree canopy in 37 US Cities. *npj Urban Sustainability* 1(1).

Maloy, Otis C. 2003. White pine blister rust. *The Plant Health Instructor* .

Michaels, Guy and Ferdinand Rauch. 2018. Resetting the urban network: 117–2012. *The Economic Journal* 128(608): 378–412.

Michney, Todd M. 2022. How the City Survey's Redlining Maps Were Made: A Closer Look at HOLC's Mortgagee Rehabilitation Division. *Journal of Planning History* 21(4): 316–344.

Miñano-Mañero, Alba. 2023. *I see trees of green*: constructing panels of tree canopy with aerial imagery .

Morales, Dominic J. 1980. The contribution of trees to residential property value. *Journal of Arboriculture* 6(11): 305–308.

Namin, S., W. Xu, Y. Zhou, and K. Beyer. 2020. The legacy of the Home Owners' Loan Corporation and the political ecology of urban trees and air pollution in the United States. *Social Science & Medicine* 246: 112758.

Nardone, Anthony, Kara E. Rudolph, Rachel Morello-Frosch, and Joan A. Casey. 2021. Redlines and greenspace: The relationship between historical redlining and 2010 greenspace across the united states. *Environmental health perspectives* 129(1): 017006.

Nelson, Robert K., LaDale Winling, Richard Marciano, Nathan Connolly *et al.* 2017. Mapping inequality. *American panorama* .

Netusil, Noelwah R., Sudip Chattopadhyay, and Kent F. Kovacs. 2010. Estimating the demand for tree canopy: a second-stage hedonic price analysis in portland, oregon. *Land Economics* 86(2): 281–293.

Neumark, David and Helen Simpson. 2015. Place-Based Policies. In Gilles Duranton, J. Vernon Henderson, and William C. Strange (eds.) *Handbook of Regional and Urban Economics*, *Handbook of Regional and Urban Economics*, volume 5. Elsevier, 1197–1287.

Nowak, David J and Tim Aevermann. 2019. Tree compensation rates: Compensating for the loss of future tree values. *Urban Forestry & Urban Greening* 41: 93–103.

Ost, Ben, Weixiang Pan, and Douglas Webber. 2018. The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies. *Journal of Labor Economics* 36(3): 779–805.

Pandit, Ram, Maksym Polyakov, and Rohan Sadler. 2014. Valuing public and private urban tree canopy cover. *Australian Journal of Agricultural and Resource Economics* 58(3): 453–470.

Rappaport, Jordan. 2007. Moving to nice weather. *Regional Science and Urban Economics* 37(3): 375–398.

Rappaport, Jordan and Jeffrey Sachs. 2003. The United States as a Coastal Nation. *Journal of Economic Growth* 8: 5–46.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and III Owens, Raymond. 2010. Housing Externalities. *Journal of Political Economy* 118(3): 485–535.

Rothstein, Richard. 2017. *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing.

Sampson, Robert J. 2016. Individual and community economic mobility in the Great Recession era: The spatial foundations of persistent inequality. In A Brown, D Buchholz, D Davis, and A González (eds.) *Economic Mobility: Research and Ideas on Strengthening Families, Communities and the Economy*. St. Louis, MO: Federal Reserve Bank of St. Louis and the Board of Governors of the Federal Reserve System, 261–286.

Sampson, Robert J. 2019. Neighbourhood effects and beyond: Explaining the paradoxes of inequality in the changing American metropolis. *Urban Studies* 56(1): 3–32.

Shertzer, Allison, Tate Twinam, and Randall P. Walsh. 2016. Race, Ethnicity, and Discriminatory Zoning. *American Economic Journal: Applied Economics* 8(3): 217–46.

Shi, Ying, Daniel Hartley, Bhash Mazumder, and Aastha Rajan. 2022. The effects of the Great Migration on urban renewal. *Journal of Public Economics* 209: 104647.

Sood, Aradhya and Kevin Ehrman-Solber. 2023. Long Shadow of Housing Discrimination: Evidence from Racial Covenants .

Tubby, K. V. and J. F. Webber. 2010. Pests and diseases threatening urban trees under a changing climate. *Forestry: An International Journal of Forest Research* 83(4): 451–459.

Twinam, Tate. 2017. Danger zone: Land use and the geography of neighborhood crime. *Journal of Urban Economics* 100: 104–119.

Twinam, Tate. 2018. The long-run impact of zoning: Institutional hysteresis and durable capital in Seattle, 1920–2015. *Regional Science and Urban Economics* 73: 155–169.

Villarreal, Carlos. 2014. Where the other half lives: Evidence on the origin and persistence of poor neighborhoods from New York City 1830-2012. *Unpublished MS. https://sites. google. com/site/carlosrvillarreal* 1(2): 3.

Voth, Hans-Joachim. 2021. Chapter 9 - persistence – myth and mystery. In Alberto Bisin and Giovanni Federico (eds.) *The Handbook of Historical Economics*. Academic Press, 243–267.

Wachter, Susan M. and Grace Wong. 2008. What is a tree worth? Green-city strategies, signaling and housing prices. *Real Estate Economics* 36(2): 213–239.

Wheelock, David C. *et al.* 2008. The Federal response to home mortgage distress: Lessons from the Great Depression. *REVIEW-FEDERAL RESERVE BANK OF SAINT LOUIS* 90(3): 133.

Wilson, Barry T., Andrew J. Lister, Rachel I. Riemann, and Douglas M. Griffith. 2013. Live tree species basal area of the contiguous United States (2000-2009) .

Winling, LaDale C and Todd M Michney. 2021. The Roots of Redlining: Academic, Governmental, and Professional Networks in the Making of the New Deal Lending Regime. *Journal of American History* 108(1): 42–69.

Woods, Louis Lee. 2012. The Federal Home Loan Bank Board, redlining, and the national proliferation of racial lending discrimination, 1921–1950. *Journal of Urban History* 38(6): 1036–1059.

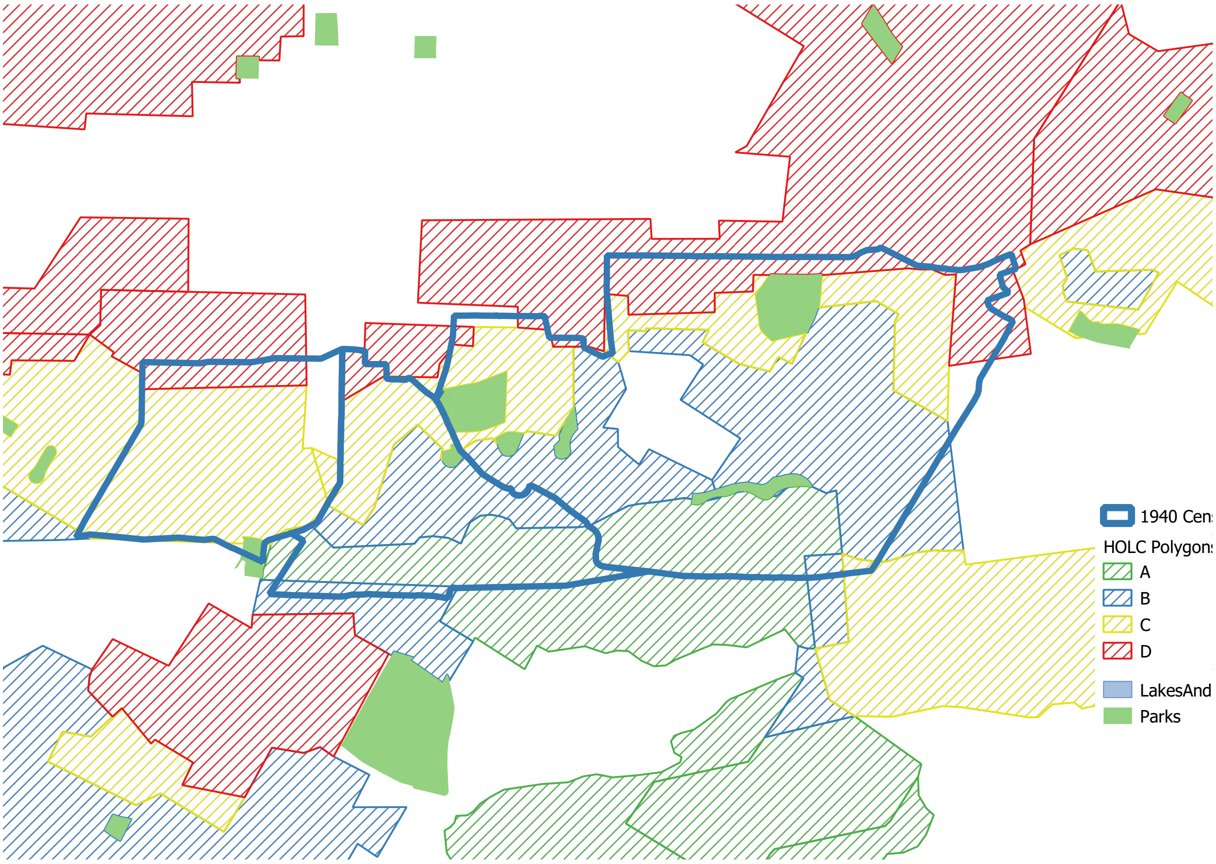
Yang, Lin, Xiaqing Wu, Emil Praun, and Xiaoxu Ma. 2009. Tree Detection from Aerial Imagery. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '09. New York, NY, USA: Association for Computing Machinery, 131–137.

Zuk, Miriam, Ariel H. Bierbaum, Karen Chapple, Karolina Gorska, and Anastasia Loukaitou-Sideris. 2018. Gentrification, Displacement, and the Role of Public Investment. *Journal of Planning Literature* 33(1): 31–44.

Appendix

8.1 Census-to-Redlining Constant Crosswalks

Figure 8.1.1: Redlining maps and 1940 tracts



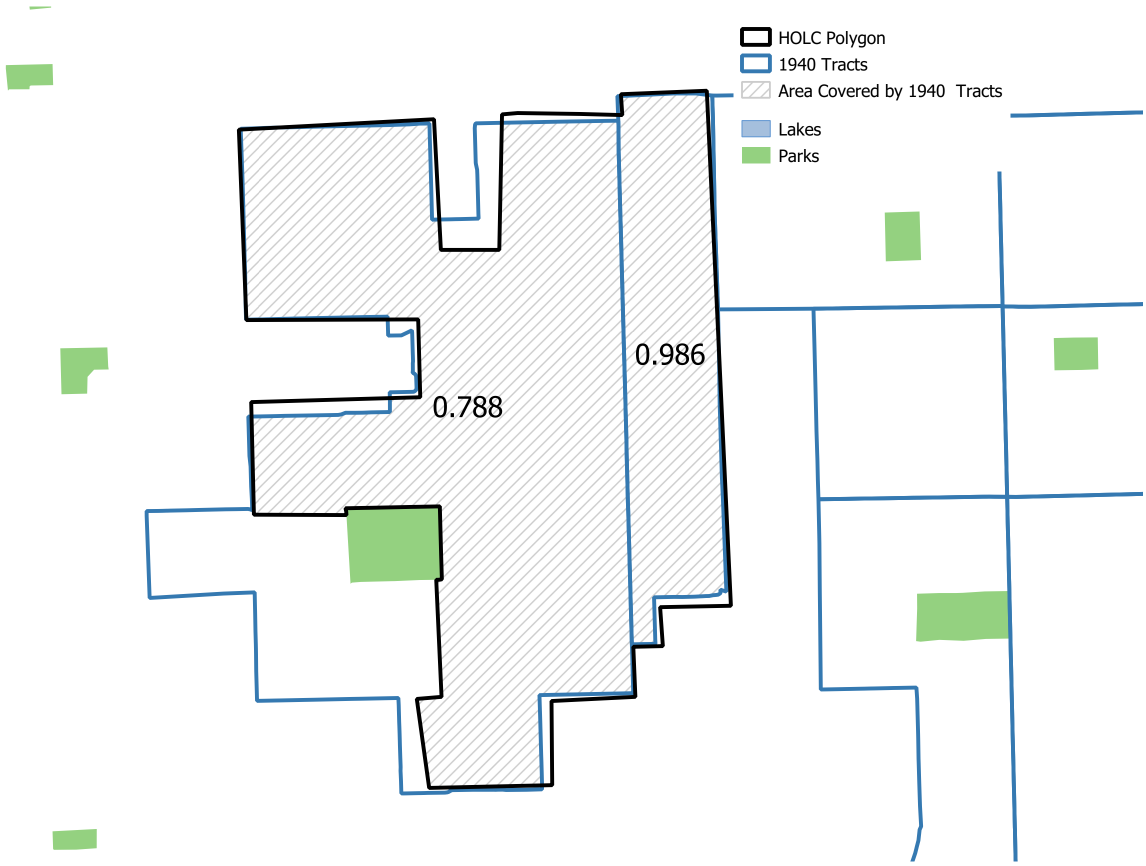
Notes: This figure shows the intersection between HOLC graded neighborhoods from the redlining maps and the 1940 Census tracts. Source: See data description. Own elaboration.

As shown in Figure 8.1.1, Census units do not align perfectly with the original neighborhoods of the redlining maps. As a result, one needs to develop a matching method to merge redlining and Census information. As discussed in the main text, assigning grades to Census units generates a series of problems that undermine the validity of the results obtained with that procedure. [44] As a result, I follow the opposite strategy and assign Census units to graded neighborhoods by using the Census-to-Redlining Constant Crosswalks, which I describe in detail in this Appendix.

The basic idea behind the crosswalks is to compute the share of the Census units that fall in the original graded neighborhood. Then, one can use these weights to construct data at the originally graded neighborhood. As shown in Figure 8.1.2, the black line represents the graded neighborhood (HOLC polygon), and the blue line the two tracts that intersect it. For one of them, 78% of its area is contained in the neighborhood. For the other, 98% of it falls in

[44]This procedure eliminates the measurement error on the grade assignment but the concern of measurement error induced on the neighborhood Census variables by the areal weights would still be present. However, when performing regressions this measurement error will not bias the results as long as it is uncorrelated with the error term.

Figure 8.1.2: Census-to-Redlining Constant Crosswalks



Notes: This figure is an example of the areal weights used in the Census-to-Redlining Constant Crosswalks. Source: see data description. Own elaboration.

the neighborhood. As a result, data at the graded neighborhood level will be the weighted sum of the data for these two tracts, with these areal weights.

Constructing the weights for 1940 is straightforward and is simply done by intersecting 1940 tracts with the originally graded neighborhoods. Then, I compute the area of the intersection. The result will be a file that contains each graded neighborhood, its area, the tracts that intersect them and their area, and the area of the intersection. From this, I first compute the share of the HOLC neighborhood that is covered by 1940 tracts and keep only the ones that are covered by at least 80%. Then, I simply compute the share of the tract that falls in the neighborhood. For the rest of the years, the process is essentially the same but it becomes more cumbersome since I need to restrict the area to the one covered in 1940. I will use 1950 tracts for the explanation for simplicity but this is the procedure applied to any other decade besides 1940. To do this, I performed the same intersection between 1950 tracts and the HOLC-graded neighborhoods. Then, I re-intersect this with the 1940-HOLC intersection. Computing the areas of these re-intersections tells the area of the 1950 tract that was covered already in 1940. Then I simply compute the share of the 1950 tract that falls in this re-intersection area. Since a 1950 tract does not necessarily intersect with only one 1940 tract, I then sum the different weights of the 1950 tract that falls in the same graded neighborhood (i.e., I am just adding the area share of the 1950 tract that corresponds to a 1940 tract and the area share of this same tract that corresponds to other 1940 tract that fall in the same graded neighborhood).

The result from applying this procedure every decade is a set of files that have four columns: the HOLC neighborhood (index), the assigned HOLC grade (A-B-C-D), the identifier for the tract/block group that falls in it (GISJOIN), and the area share of the tract/block group that corresponds to that Census unit-HOLC neighborhood intersection. Then, to construct data at the neighborhood level, one only needs to download data from the National Register of Historical Places at the tract level (1940-1980) and block group level (1990-2015). The use of the crosswalks is essentially the same as the use of Lee and Lin (2018)'s ones. Some examples below illustrate how the variables in this paper have been constructed.

1. Merge the Census data with the Census-to-Redlining Constant Crosswalks

```
import delimited "$data\NHGIS_1940.csv", rowrange(1:) varnames(1) clear
rename gisjoin gisjoin1940
merge 1:m gisjoin1940 using "$cw\HOLCto1940.dta" /*Crosswalk file*/
keep if _merge == 3
drop _merge
```

2. For variables expressed as counts, simply weight them with the areal weight `ch1940` (share of the 1940 tract that falls in the HOLC graded neighborhood)

```
local varlist "white population"
foreach var of local varlist{
gen wt‘var’ = ‘var’*ch1940
}
```

3. For variables expressed as counts, add the weighted observations of the previous step at the HOLC neighborhood level

```
local varlist "white population"
foreach var of local varlist{
bysort index: egen ‘var’1940 = total(wt‘var’), missing
}
```

4. Generate the variable of interest, drop duplicates and save

```
gen WhiteShare1940 = white1940/population1940
keep index WhiteShare1940 population1940
egen tag = tag(index)
keep if tag
save "$data\population1940.dta", replace
```

5. For home values and income, after obtaining the MSA medians, apply the crosswalks to the number of housing units or families in each interval, attach the midpoint of the interval or the value if it is the first or last reported interval and compute the share on and above the MSA median for each graded neighborhood.
6. For variables reported as means, apply the crosswalks to the relevant count variable, attach the value, and obtain the mean in the polygon. For instance, for average family income obtain the cross-walked number of families, multiply it by the reported income, and average it for each neighborhood.

8.2 *Waterfront modifications data sources*

This appendix describes the waterfront modifications that were considered together with their data source. Details on reasons why some cities are not considered here as well as additional information can be made available upon request. These modifications have been merged into a single file that is also available upon request. [45]

Baltimore: Data for waterfront improvements comes directly from the digitized Urban Renewal Plans of the Baltimore Department of Planning. The date for the Canton Waterfront comes from the same plan, which was approved in 1984. By establishing the date to be 1990, the modification taking into account the date will only appear from 2000 onward, giving a long enough time window for it to have taken place. For Inner Harbor, the project was approved in 1967, the date chosen is because in 1976 a series of celebrations of the US Bicentennial took place there, suggesting the project had already been, at least partially, completed.

Boston: Boston Waterfront modifications include the creation of the Cristopher Columbus Waterfront Park and the Harborwalk. The location of the first one comes from selecting that park from the Open Space shapefile provided by the City of Boston Open data portal. The area is meant to capture the waterfront and the redevelopment of the Faneuil Hall area. The Harborwalk is obtained by extracting it from the shapefile containing shared walker trails from the following dataset. Dates are based on the New York's Time Article *"BOSTON WATERFRONT: AT 25, A MODEL URBAN RENEWAL"* (1986) available here.

Bronx: All data comes from the New York City Departments of Parks and & Recreation. The parks are extracted from the Open Space Recreation Parks shapefile provided by the City of New York. The Bronx River Greenway is obtained by merging the Bronx Park and the Shoelace since no park with such a name appeared on the shapefile.

Brooklyn: Data for the Brooklyn Bridge Park is obtained in the same way as the data for the Bronx. Only the completed parts of the Brooklyn Waterfront Greenway are considered. They are obtained by extracting the objects designed as greenways from the New York Biking Routes shapefile and comparing them to the ones provided by the Brooklyn Greenway Initiative (BGI). The attached date is based on the information given by the BGI.

Buffalo: Although the waterfront redevelopment of Buffalo is not considered in the analysis because the buffer around it does not intersect any graded neighborhood, the modification considered is the redevelopment of Canalside. It was geolocated with the coordinates of Canalside on Google Maps. The attached date was 2008 when the Central Wharf was inaugurated. More information can be found here.

Cambridge: The modifications considered come from the Cambridge Community Development Department. It considers the 1978 East Cambridge Riverfront Plan and the 1983 Cambridgeport Revitalization Plan. It was geolocated by extracting the districts of East Waterfront and Cambridgeport.

Chicago: Modifications considered include the Riverwalk and the Lake Front Trail. The information on the Riverwalk was obtained from the Chicago River Timeline from the Chicago River Edge Ideas Lab which depends on the City of Chicago's Department of Planning and

[45] For some areas, it was unclear whether a modification had taken place or not and there were incongruities among data sources. As a result, I only considered the modifications that according to the majority of data sources had been fully implemented and, in case of doubt, by inspecting the area in Google Street View and comparing it to the rest of the areas before deciding.

Development. It was geolocated with the Open Spaces-Riverwalk shapefile of the Chicago Data Portal. The date was chosen because it was when the construction between Lake Shore Drive and Michigan Avenue started. Data on the Lake Front Trail is extracted similarly from the Bike Routes of the Chicago Data Portal. It was designated as a bike trail in 1963.

Columbus: The considered modifications were extracted because of their appearance in the case study "The transformation of the downtown Columbus riverfront 1998-2020" by the City of Columbus and MKSK studios, which can be accessed here. The created parks (Genoa Park, Lower Scioto Park, and North Bank Park) were extracted from the City of Columbus Open Data Park Property Boundaries shapefile.

Duluth: only considers the Canal Park. It was chosen because of the Duluth New Tribune 2010 Article *"History: Changing Duluth's waterfront from junk to jewel of the North"*, accessible here. It was geolocated by extracting all addressed structures in Canal Park from the Address Point shapefile of the St. Luois County (MN) data portal.

Indianapolis: information on Canal Walk was obtained from the Cultural Landscape Foundation. It was geolocated by extracting the objects named Canal Walk from the Indianapolis Parks shapefile provided by the City of Indianapolis data portal.

Louisville: Information for the Waterfront Park was obtained from its web page. It was located by extracting the areas named Waterfront Park from the Louisville Metro Areas of Interest shapefile of the Louisville Open Geospatial Data portal.

Lower Westchester County: The only modification is a part of the Bronx River Parkway that intersects one neighborhood there. See the description for the Bronx.

Manhattan: Considered parks were extracted following the New York City Comprehensive Waterfront Plan (1992) and the Vision 2020: New York City Comprehensive Waterfront Park (2011). They are all extracted from the Open Space shapefile of the NYC data portal.

Minneapolis: Nicollet Island was deemed as a modification following this newspaper article. Even if other areas could have been relevant (i.e., Hennepin Island, Promenade Main Street, West Bank Waterfront, Basett's Creek) I was only able to locate Nicollet Island by extracting the parks with such names from Minneapolis Open Data. Moreover, with Google Street View these areas, as well as the riverbank, did not seem to have been developed comparably to other areas in other cities.

New Orleans: Although it does not intersect any neighborhood, the modifications considered were the ones that took place around the French Quarters (Moonwalk and Woldbenger Park). They were located by extracting them from the Parks data of New Orleans.

Philadelphia: Penn's Landing was considered because of the mentions in Visit Philly tourism web page. It was located by extracting the parks that would correspond to its location according to Google Maps, which would include the Irish Memorial, the Korean War Veteran's Memorial, and the Vietnam Memorial. The date was chosen since it was the inauguration of Penn's Landing Great Plaza.

Pittsburgh: The parks located are the ones that belong to the Three Rivers Parks (Monongahela, Allegheny, and Ohio) following the Pittsburgh Waterfront Master Plan. The dates and specific parks were extracted from the Pittsburgh nonprofit organization Riverlife. Besides the ones in Appendix Table 8.3.3, the Point State Park and the Northshore Riverfront Park were also considered.

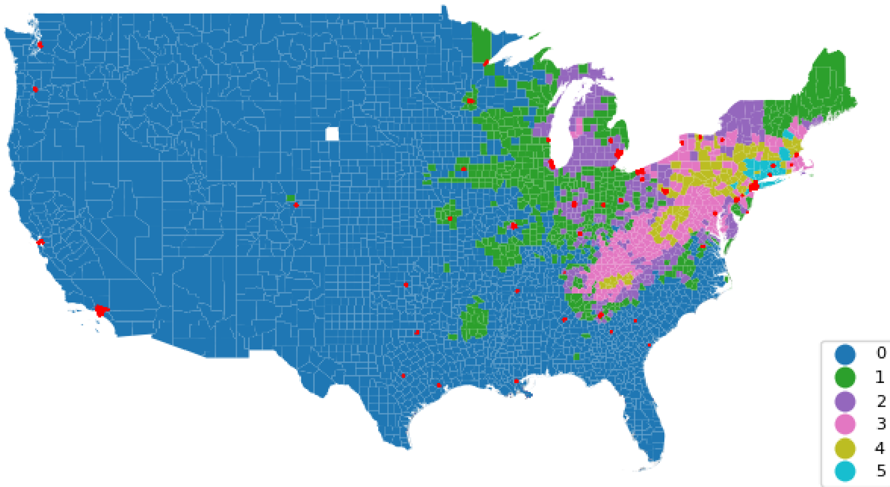
Portland: Following Portland's Park and Recreation Department, the only two considered features were the South Waterfront Park, which includes the Gov. Tom McCall Waterfront Park, and the Vera Katz Eastbank Esplanade. They were located by extracting these features from park shapefiles.

Seattle: The modifications considered to capture the Seattle waterfront redevelopment were the location of the Aquarium and the Waterfront Park.

Queens: The sole modification is a part of the Brooklyn Bridge Park that intersects neighborhoods in Queens. See Brooklyn.

8.3 Additional evidence and results

Figure 8.3.1: County distribution of pests



Notes: this map shows the total number of selected deadly plagues from (Fei, Morin, Oswalt, and Liebhold, 2019) in that county as of 2019. Source: Fei, Morin, Oswalt, and Liebhold (2019). Own elaboration.

Figure 8.3.2: Potential Emerald Ash Borer Hosts in Chicago (per thousand tree pixels)

Notes: this map shows the potential Emerald Ash Borer hosts per thousand tree pixels in Chicago HOLC neighborhoods. Source: Wilson, Lister, Riemann, and Griffith (2013) and see data description. Own elaboration.

Table 8.3.1: HOLC cities and 2010 MSA assignment

MSA 2010	HOLC City	Neighborhoods	%
Birmingham-Hoover, AL	Birmingham	310	0.8%
Los Angeles-Long Beach-Santa Ana, CA	Los Angeles	4,420	11.7%
San Francisco-Oakland-Fremont, CA	Oakland	1,420	3.8%
San Francisco-Oakland-Fremont, CA	San Francisco	1,050	2.8%
Denver-Aurora-Broomfield, CO	Denver	530	1.4%
New Haven-Milford, CT	New Haven	260	0.7%
Atlanta-Sandy Springs-Marietta, GA	Atlanta	1,210	3.2%
Augusta-Richmond County, GA-SC	Augusta	260	0.7%
Macon, GA	Macon	410	1.1%
Chicago-Joliet-Naperville, IL-IN-WI	Chicago	3,360	8.9%
St. Louis, MO-IL	East St. Louis	360	1.0%
Indianapolis-Carmel, IN	Indianapolis	880	2.3%
Louisville/Jefferson County, KY-IN	Louisville	510	1.3%
New Orleans-Metairie-Kenner, LA	New Orleans	1,190	3.1%
Boston-Cambridge-Quincy, MA-NH	Boston	390	1.0%
Boston-Cambridge-Quincy, MA-NH	Cambridge	150	0.4%
Boston-Cambridge-Quincy, MA-NH	Somerville	10	0.0%
Baltimore-Towson, MD	Baltimore	450	1.2%
Detroit-Warren-Livonia, MI	Detroit	2,330	6.2%
Flint, MI	Flint	530	1.4%
Duluth, MN-WI	Duluth	340	0.9%
Minneapolis-St. Paul-Bloomington, MN-WI	Minneapolis	860	2.3%
Kansas City, MO-KS	Greater Kansas City	520	1.4%
St. Louis, MO-IL	St. Louis	1,370	3.6%
Atlantic City-Hammonton, NJ	Atlantic City	70	0.2%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	Camden	200	0.5%
Trenton-Ewing, NJ	Trenton	80	0.2%
New York-Northern New Jersey-Long Island, NY-NJ-PA	Bronx	450	1.2%
New York-Northern New Jersey-Long Island, NY-NJ-PA	Brooklyn	670	1.8%
Buffalo-Niagara Falls, NY	Buffalo	380	1.0%
New York-Northern New Jersey-Long Island, NY-NJ-PA	Lower Westchester Co.	480	1.3%
New York-Northern New Jersey-Long Island, NY-NJ-PA	Manhattan	530	1.4%
New York-Northern New Jersey-Long Island, NY-NJ-PA	Queens	1,770	4.7%
Rochester, NY	Rochester	320	0.8%
New York-Northern New Jersey-Long Island, NY-NJ-PA	Staten Island	730	1.9%
Syracuse, NY	Syracuse	420	1.1%
Akron, OH	Akron	550	1.5%
Cleveland-Elyria-Mentor, OH	Cleveland	1,960	5.2%
Columbus, OH	Columbus	600	1.6%
Dayton, OH	Dayton	440	1.2%
Toledo, OH	Toledo	390	1.0%
Portland-Vancouver-Hillsboro, OR-WA	Portland	1,010	2.7%
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	Philadelphia	740	2.0%
Pittsburgh, PA	Pittsburgh	1,100	2.9%
Nashville-Davidson–Murfreesboro–Franklin, TN	Nashville	10	0.0%
Dallas-Fort Worth-Arlington, TX	Dallas	260	0.7%
Richmond, VA	Richmond	430	1.1%
Seattle-Tacoma-Bellevue, WA	Seattle	600	1.6%
Milwaukee-Waukesha-West Allis, WI	Milwaukee Co.	480	1.3%

Notes: this table displays the MSA-HOLC city assignment, together with the amount of neighborhoods in each city for the entire 1940-2015 period. Source: see data description. Own elaboration.

Table 8.3.2: Distribution of neighborhoods per grade

HOLC City	*A-graded (Green)*	*B-graded (Blue)*	*C-graded (Yellow)*	*D-graded (Red)*	Total
Akron	90	170	180	110	550
Atlanta	100	300	450	360	1,210
Atlantic City	0	10	50	10	70
Augusta	0	60	70	130	260
Baltimore	50	140	150	110	450
Birmingham	0	90	150	70	310
Boston	10	80	180	120	390
Bronx	20	120	230	80	450
Brooklyn	10	170	250	240	670
Buffalo	50	120	120	90	380
Cambridge	10	70	50	20	150
Camden	10	30	80	80	200
Chicago	70	560	1,760	970	3,360
Cleveland	350	550	780	280	1,960
Columbus	50	220	230	100	600
Dallas	60	90	60	50	260
Dayton	40	80	130	190	440
Denver	60	130	190	150	530
Detroit	150	390	1,180	610	2,330
Duluth	40	70	130	100	340
East St. Louis	40	50	120	150	360
Flint	20	70	180	260	530
Greater Kansas City	30	110	190	190	520
Indianapolis	50	190	280	360	880
Los Angeles	600	1,220	1,800	800	4,420
Louisville	80	150	160	120	510
Lower Westchester Co.	80	70	210	120	480
Macon	20	60	160	170	410
Manhattan	80	120	60	270	530
Milwaukee Co.	30	110	210	130	480
Minneapolis	180	280	230	170	860
Nashville	0	0	10	0	10
New Haven	20	40	120	80	260
New Orleans	80	180	440	490	1,190
Oakland	120	460	570	270	1,420
Philadelphia	70	240	180	250	740
Pittsburgh	110	270	410	310	1,100
Portland	110	320	450	130	1,010
Queens	10	180	1,130	450	1,770
Richmond	20	90	90	230	430
Rochester	20	70	160	70	320
San Francisco	130	370	360	190	1,050
Seattle	130	180	180	110	600
Somerville	0	0	10	0	10
St. Louis	320	440	450	160	1,370
Staten Island	40	140	270	280	730
Syracuse	50	120	160	90	420
Toledo	70	120	130	70	390
Trenton	10	10	20	40	80
Total	3,690	9,110	15,160	9,830	37,790

Notes: this table shows the distribution of neighborhoods by grade-city. Source: see data description. Own elaboration.

Table 8.3.3: Geolocated waterfront modifications

HOLC City	Name geolocated modification	Date
Baltimore	Key Highway	2011
Baltimore	Middle Branch	1983
Baltimore	Canton Waterfront	1990
Baltimore	Inner Harbor East	1976
Baltimore	Fells Point Waterfront	2006
Baltimore	Inner Harbor Project I	1976
Boston	Harborwalk	1984
Boston	Christopher Columbus Park	1976
Bronx	Starlight Park	2013
Bronx	Concrete Plant Park	2009
Bronx	Soundview Park	1998
Bronx	Bronx River Parkway	2000
Bronx	Bronx Park	2000
Bronx	Hunts Point Riverside Park	2007
Brooklyn	Brooklyn Greenway	2010
Brooklyn	Brooklyn Bridge Park	2010
Brooklyn	Greenpoint-Williamsburg Waterfront	2005
Buffalo	Canalside	2008
Cambridge	East Cambridge	1978
Cambridge	Cambridgeport	1983
Chicago	Riverwalk	2001
Chicago	Lakefront Trail	1963
Columbus	North Bank Park	2005
Columbus	Genoa Park	1999
Columbus	Lower Scioto Park	2009
Duluth	Canal Park	1993
Indianapolis	Canal Walk	2001
Louisville	Waterfront Park	1999
Manhattan	Greenway	1999
Manhattan	Riverside Park	2001
Minneapolis	Nicolette Island	1983
New Orleans	Woldenberg Park	1984
Philadelphia	Penn's Landing	1986
Pittsburgh	Point State Park	2000
Pittsburgh	Southside Riverfront Park	2012
Pittsburgh	Washington's Landing Park	1980
Pittsburgh	Northshore Riverfront Park	2001
Pittsburgh	Monongahela Wharf Landing Park	2009
Pittsburgh	Allegheny Riverfront Park	2000
Pittsburgh	Allegheny Landing Park	2000
Portland	Vera Katz Eastbank Esplanade	2000
Portland	Gov Tom McCall Waterfront Park	1978
Portland	South Waterfront Park	2000
Seattle	Seattle Aquarium	1977
Seattle	Waterfront Park	1977

Notes: This table shows the geolocated modifications, their date and the corresponding HOLC city. Source: see data description and Appendix 8.2. Own elaboration.

Table 8.3.4: HOLC cities and NAIP imagery

2010 MSA/CBSA	HOLC City	First NAIP image year	Second NAIP image year
Akron, OH	Akron	2004	2015
Atlanta-Sandy Springs-Marietta, GA	Atlanta	2007	2015
Baltimore-Towson, MD	Baltimore	2005	2015
Birmingham-Hoover, AL	Birmingham	2006	2015
Boston-Cambridge-Quincy, MA-NH	Boston	2003	2014
Boston-Cambridge-Quincy, MA-NH	Cambridge	2003	2014
Boston-Cambridge-Quincy, MA-NH	Somerville	2003	2014
Buffalo-Niagara Falls, NY	Buffalo	2006	2015
Chicago-Joliet-Naperville, IL-IN-WI	Chicago	2007	2015
Cleveland-Elyria-Mentor, OH	Cleveland	2004	2015
Columbus, OH	Columbus	2004	2015
Dayton, OH	Dayton	2004	2015
Detroit-Warren-Livonia, MI	Detroit	2005	2014
Flint, MI	Flint	2005	2014
Kansas City, MO-KS	Greater Kansas City	2007	2015
Los Angeles-Long Beach-Santa Ana, CA	Los Angeles	2005	2014
Milwaukee-Waukesha-West Allis, WI	Milwaukee Co.	2005	2015
Nashville-Davidson–Murfreesboro–Franklin, TN	Nashville	2006	2014
New Haven-Milford, CT	New Haven	2006	2014
New Orleans-Metairie-Kenner, LA	New Orleans	2007	2015
New York-Northern New Jersey-Long Island, NY-NJ-PA	Bronx	2006	2015
New York-Northern New Jersey-Long Island, NY-NJ-PA	Brooklyn	2006	2015
New York-Northern New Jersey-Long Island, NY-NJ-PA	Lower Westchester Co.	2006	2015
New York-Northern New Jersey-Long Island, NY-NJ-PA	Manhattan	2006	2015
New York-Northern New Jersey-Long Island, NY-NJ-PA	Queens	2006	2015
New York-Northern New Jersey-Long Island, NY-NJ-PA	Staten Island	2006	2015
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	Camden	2006	2015
Richmond, VA	Richmond	2003	2015
Rochester, NY	Rochester	2006	2015
San Francisco-Oakland-Fremont, CA	Oakland	2005	2014
San Francisco-Oakland-Fremont, CA	San Francisco	2005	2014
Seattle-Tacoma-Bellevue, WA	Seattle	2006	2015
St. Louis, MO-IL	East St. Louis	2007	2015
St. Louis, MO-IL	St.Louis	2007	2015
Syracuse, NY	Syracuse	2006	2015
Toledo, OH	Toledo	2004	2015
Trenton-Ewing, NJ	Trenton	2006	2015

Notes: this table shows the cities with available NAIP imagery and the two years years considered to predict tree canopy. Source: see data description and Appendix. Own elaboration. par

Table 8.3.5: Distribution of population in 1940 by HOLC grade

	% population	% white	% black
A-graded (Green)	3%	3%	1%
B-graded (Blue)	16%	18%	2%
C-graded (Yellow)	41%	44%	10%
D-graded (Red)	40%	35%	87%

Notes: this table shows the distribution of population in 1940 per grade, for a given decade, for neighborhoods with Census data. Source: see data description. Own elaboration.

Table 8.3.6: Distribution water & park amenities by HOLC grade

	No water & parks amenities	Water & park amenities	Total
A-graded (Green)	132	237	369
B-graded (Blue)	339	572	911
C-graded (Yellow)	604	912	1,516
D-graded (Red)	365	618	983
Total	1,440	2,339	3,779

*Notes:*this table shows the distribution of water and park amenities, for a given decade, for neighborhoods with Census data. Source: see data description. Own elaboration.

Table 8.3.7: Distribution waterfront modifications by HOLC grade

	No waterfront modifications	Waterfront modification	Total water & park amenities
A-graded (Green)	82	5	237
B-graded (Blue)	141	23	572
C-graded (Yellow)	215	25	912
D-graded (Red)	188	36	618
Total	626	89	2,339

Notes: this table shows the distribution of waterfront modifications, for a given decade, for neighborhoods with Census data. Source: see data description. Own elaboration

Table 8.3.8: Descriptive statistics, 1940

	% white	% housing units above MSA median home value
A-graded (Green)		
Mean	98%	89%
Std. Dev.	5	12
B-graded (Blue)		
Mean	98%	79%
Std. Dev.	05	16
C-graded (Yellow)		
Mean	97%	63%
Std. Dev.	8	20
D-graded (Red)		
Mean	86%	43%
Std. Dev.	22	21
Total		
Mean	94%	64%
Std. Dev.	14	24

Notes: this table shows the descriptive statistics of the relevant variables for 1940. Source: see data description. Own elaboration.

Table 8.3.9: Descriptive statistics, 1950

	% white	% housing units above MSA median home value	% families above MSA median family income
A-graded (Green)			
Mean	98%	91%	70%
Std. Dev.	3	13	10
B-graded (Blue)			
Mean	98%	81%	68%
Std. Dev.	5	17	11
C-graded (Yellow)			
Mean	96%	64%	63%
Std. Dev.	9	23	12
D-graded (Red)			
Mean	82%	41%	51%
Std. Dev.	26	25	15
Total			
Mean	93%	0.65%	0.62%
Std. Dev.	16	27	14

Notes: this table shows the descriptive statistics of the relevant variables for 1950. Source: see data description. Own elaboration.

Table 8.3.10: Descriptive statistics, 2015

	% white	% housing units above MSA median home value	% families above MSA median family income
A-graded (Green)			
Mean	72%	75%	72%
Std. Dev.	26	28	18
B-graded (Blue)			
Mean	62%	57%	56%
Std. Dev.	30	33	22
C-graded (Yellow)			
Mean	51%	45%	42%
Std. Dev.	31	31	21
D-graded (Red)			
Mean	44%	39%	34%
Std. Dev.	29	30	21
Total			
Mean	54%	49%	46%
Std. Dev.	31	33	24

Notes: this table shows the descriptive statistics of the relevant variables for 2015. Source: see data description. Own elaboration.

Table 8.3.11: Neighborhood change: 1950-2015

Panel A: % neighborhoods below* msa *average in 1950			
	White %	% housing units above MSA median home value	% families above MSA median family income
A-graded (Green)	5%	4%	18%
B-graded (Blue)	6%	16%	20%
C-graded (Yellow)	15%	48%	40%
D-graded (Red)	47%	79%	72%

Panel B: % neighborhoods remaining still below* msa *average in 2015			
	White share	% housing units above MSA median home value	% families above MSA median family income
A-graded (Green)	16%	69%	3%
B-graded (Blue)	54%	63%	31%
C-graded (Yellow)	75%	70%	65%
D-graded (Red)	79%	69%	75%

Notes: this table shows the share of HOLC neighborhoods below the MSA means in 1950. and the share still below in 2015. Source: see data description. Own elaboration.

Table 8.3.12: Descriptive statistics for tree pixels

	% tree pixels 2000	% tree pixels 2015	Tree growth
A-graded (Green)			
Mean	30%	37%	112%
Std. Dev.	25	26	307
B-graded (Blue)			
Mean	23%	29%	225%
Std. Dev.	25	024	906
C-graded (Yellow)			
Mean	18%	23%	316%
Std. Dev.	21	20	709
D-graded (Red)			
Mean	16%	21%	381%
Std. Dev.	20	20	1551
Total			
Mean	20%	25%	292%
Std. Dev.	23	22	1015

Notes: this table shows the descriptive statistics for the share of tree pixels and tree growth. Source: see data description. Own elaboration.

Table 8.3.13: Induced measurement error, population counts

	(1)	(2)	(3)
Geographic unit:	HOLC neighborhood	Census-to-Redlining crosswalk	Redlining-to-Census crosswalk
D-graded	-2,541.61***	-2,216.91***	-290.84***
	(565)	(530)	(78)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	8669.24	8017.92	3254.48
Observations	1,350	1,350	2,760
Adjusted R^2	0.43	0.45	0.37
Adjusted within R^2	0.03	0.02	0.02

Notes: All columns contain border-pair, so all coefficients are estimated on the basis of within D-C pairs. The dependent variable is population counts 2010 in each geographic unit obtained from raster data. Samples are restricted to adjacent D-C neighborhoods. Column (1) represents the results for the true values in HOLC polygons; Column (2) employs the Census-to-Redlining Crosswalks to population counts obtained for 2010 tracts to perform the regression at the HOLC level; Column (3) estimates the regression at the 2010 tract level, assigning grades to tracts based on spatial overlap. Standard errors are clustered by Census division level and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.14: Induced measurement error, population density

	(1)	(2)	(3)
Geographic unit	HOLC neighborhood	Census-to-Redlining crosswalk	Redlining-to-Census crosswalk
D-graded	-422.02***	-400.22**	-240.06***
	(41.52)	(148.93)	(58.14)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	3225.11	2899.59	5366.61
Observations	1,350	1,350	2,760
Adjusted R^2	0.80	0.88	0.74
Adjusted within R^2	0.03	0.06	0.00

Notes: All columns contain border-pair, so all coefficients are estimated on the basis of within D-C pairs. The dependent variable is population density (per square kilometer) in each geographic unit obtained from raster data in 2010. Samples are restricted to adjacent D-C neighborhoods. Column (1) represents the results for the true values in HOLC polygons directly obtained from raster data; Column (2) employs the Census-to-Redlining Crosswalks to population counts obtained from raster data directly at the 2010 tracts to perform the regression at the HOLC level; Column (3) estimates the regression at the 2010 tract level, assigning grades to tracts based on spatial overlap. Standard errors are clustered by Census division level and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.15: Persistence of redlining assigning grades to 1940 tracts, all D-C tracts

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-18.67***	-22.26***	-16.58***
	(1.11)	(1.74)	(0.77)
D-graded × Post^{1977}	9.25***	18.66***	8.99***
	(1.38)	(2.19)	(1.12)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	60.99	37.5	36.97
Observations	43,399	42,411	38,495
Adjusted R^2	0.33	0.24	0.25
Adjusted within R^2	0.05	0.08	0.10
Average Persistence	50%	16%	46%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. The sample consists of all 1940 tracts assigned a D-C grade. The grade assignment is based on the spatial overlap between grades and 1940 Census tracts. The *Post*^{1977} period is 1980-2015. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before. Due to data availability, columns (1) and (2) are estimated for the 1940-1980 period and column (3) for 1950 -1980. Standard errors are clustered by Census division- decade and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.16: Persistence of redlining grades to 1940 tracts, bordering D-C tracts

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-12.12***	-15.14***	-7.96***
	(1.05)	(1.08)	(0.52)
D-graded × Post^{1977}	6.83***	10.67***	3.13***
	(1.23)	(1.29)	(0.64)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	60.74	35.38	35.83
Observations	19,780	19,539	17,570
Adjusted R^2	0.70	0.48	0.57
Adjusted within R^2	0.05	0.07	0.06
Average Persistence	44%	29%	61%

Notes: All columns contain border-pair and year fixed effects, so coefficients are estimated on the basis of within D-C pairs. The sample consists of adjacent 1940 tracts assigned a D-C grade that share the longest border. The grade assignment is based on the spatial overlap between grades and 1940 Census tracts. The *Post*^{1977} period is 1980-2015. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before. Due to data availability, columns (1) and (2) are estimated for the 1940-1980 period and column (3) for 1950 -1980. Standard errors are clustered by Census division- decade and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.17: Within MSA persistence of redlining, 1980

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-13.11***	-18.92***	-11.67***
	(1.08)	(1.33)	(0.64)
D-graded × Post^{1977}	0.41	5.73***	2.50***
	(1.30)	(1.86)	(0.85)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	66.36	44.05	42.82
Observations	12,423	12,189	9,911
Adjusted R^2	0.30	0.26	0.34
Adjusted within R^2	0.06	0.12	0.13
Average Persistence	97%	70%	79%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. The *Post*^{1977} period is restricted to 1980. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before. Due to data availability, columns (1) and (2) are estimated for the 1940-1980 period and column (3) for 1950 -1980. Standard errors are clustered by Census division - year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.18: Within bordering D-C neighborhoods persistence of redlining, 1980

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-8.23***	-10.75***	-6.13***
	(0.66)	(0.94)	(0.46)
D-graded × Post^{1977}	2.86	5.83***	1.98**
	(1.97)	(2.08)	(0.86)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	62.36	38.98	39.01
Observations	6,110	6,005	4,878
Adjusted R^2	0.66	0.63	0.73
Adjusted within R^2	0.05	0.10	0.12
Average Persistence	65%	46%	68%

Notes: All columns contain border-pair and year fixed effects, so all coefficients are estimated on the basis of within D-C pair. The *Post*^{1977} period is restricted to 1980. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before. Due to data availability, columns (1) and (2) are estimated for the 1940-1980 period and column (3) for 1950 -1980. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.19: Within placebo D-C pair

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income income
Placebo D-graded	0.436 (1.1201)	0.055 (0.5086)	0.224 (0.4351)
Water or park amenities	-1.763 (1.1120)	-2.086** (0.9905)	-2.091*** (0.6739)
Placebo D-graded × Water or park amenities	-0.063 (1.4098)	1.351* (0.7626)	0.465 (0.5311)
Placebo D-graded × Post^{1977}	0.277 (1.3250)	0.321 (0.8808)	-0.016 (0.6192)
Water or park amenities × Post^{1977}	5.119*** (1.1404)	2.496* (1.4049)	1.728* (0.8918)
Placebo D-graded × Water or park amenities × Post^{1977}	-0.208 (1.6185)	-1.186 (1.1733)	-0.063 (0.7193)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	71.05	53.96	49.23
Observations	8,893	8,790	7,902
Adjusted R^2	0.71	0.58	0.66
Adjusted within R^2	0.00	0.00	0.00

Notes: All columns contain border-pair and year fixed effects, so all coefficients are estimated on the basis of within placebo D-C pair. The placebo D-C pairs are found after assigning the placebo grades to all neighborhoods by keeping the pair that shares the longest border with the placebo D-graded and is longer than 500 meters. The $Post^{1977}$ period is 1980-2015. Water or park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 20% of the area. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.20: All D-C neighborhoods and water or parks above MSA median

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-14.501***	-14.257***	-10.178***
	(1.4003)	(1.5557)	(0.8313)
D-graded × Post^{1977}	5.126***	7.165***	2.371**
	(1.6494)	(2.0016)	(1.0502)
Water or park amenities	1.878**	1.394	0.415
	(0.7910)	(1.0148)	(0.5615)
Water or park amenities × Post^{1977}	5.790***	1.991	2.201**
	(1.1782)	(1.5897)	(0.9423)
D-graded × Water or park amenities	1.537	-6.884***	-1.963**
	(1.0603)	(0.8685)	(0.7785)
D-graded × Water or park amenities × Post^{1977}	-2.442*	6.002***	0.971
	(1.3277)	(1.4832)	(1.1797)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.28
Adjusted within R^2	0.05	0.07	0.08
Average Persistence Water or Parks	79%	38%	72%
Average Persistence No Water nor Parks	65%	50%	77%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. The *Post*^{1977} period is 1980-2015. Water or park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least a share of the neighborhood area larger than the MSA median share of any of the features. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without water or parks. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.21: Bordering D-C neighborhoods and water or parks above MSA median

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-8.04***	-7.98***	-5.46***
	(1.03)	(1.16)	(0.64)
Water or park amenities	-1.18	-1.11	-1.22*
	(1.20)	(1.00)	(0.64)
D-graded × Water or park amenities	-0.28	-4.33***	-1.02
	(1.23)	(0.94)	(0.73)
D-graded × $Post^{1977}$	4.48***	5.40***	2.13**
	(1.53)	(1.47)	(0.89)
Water or park amenities × $Post^{1977}$	6.12***	5.06***	3.94***
	(1.44)	(1.54)	(0.84)
D-graded × Water or park amenities × $Post^{1977}$	-1.11	2.91*	-0.76
	(1.60)	(1.62)	(1.09)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	62.36	38.98	39.01
Observations	11,030	10,925	9,798
Adjusted R^2	0.73	0.54	0.63
Adjusted within R^2	0.04	0.05	0.06
Average Persistence Water or Parks	59%	32%	79%
Average Persistence No Water nor Parks	44%	32%	61%

Notes: All columns contain border-pair and year fixed effects, so all coefficients are estimated on the basis of within placebo D-C pair. Water and park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least a share of the neighborhood area larger than the MSA median share of any of the features. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without amenities. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.22: Natural amenities (10% threshold) mitigate the persistence of redlining, all D-C neighborhoods within the same MSA

	(1)	(2)	(3)
Dependent variables	% white	% housing units above MSA median home value	% families above MSA median family income
D-graded	-13.15***	-14.23***	-9.53***
	(1.08)	(1.58)	(0.94)
D-graded × Post^{1977}	3.54**	5.90***	1.15
	(1.52)	(2.18)	(1.16)
Water or park amenities	2.37***	1.65	0.66
	(0.64)	(1.04)	(0.51)
Water or park amenities × Post^{1977}	3.89***	1.06	1.77*
	(1.09)	(1.54)	(0.89)
D-graded × Water or park amenities	-0.43	-6.34***	-2.71***
	(0.92)	(0.83)	(0.64)
D-graded × Water or park amenities × Post^{1977}	0.29	7.43***	2.71**
	(1.37)	(1.44)	(1.06)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.28
Adjusted within R^2	0.05	0.07	0.08
Average Persistence Water or Parks	72%	35%	68%
Average Persistence No Water nor Parks	73%	59%	88%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. The *Post*^{1977} period is 1980-2015. Water or park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 10% of the area. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without amenities. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.23: Natural amenities (10% threshold) mitigate the persistence of redlining, bordering D-C neighborhoods

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-7.25***	-8.12***	-5.06***
	(1.12)	(1.17)	(0.66)
D-graded × Post^{1977}	1.30	4.05**	0.20
	(1.54)	(1.55)	(1.02)
Water or park amenities	0.17	-1.08	-1.86**
	(1.04)	(1.32)	(0.84)
Water or park amenities × Post^{1977}	1.84	3.24*	2.46**
	(1.37)	(1.73)	(1.09)
D-graded × Water or park amenities	-1.43	-3.91***	-1.70***
	(1.27)	(0.90)	(0.62)
D-graded × Water or park amenities × Post^{1977}	3.91**	5.03***	2.38**
	(1.54)	(1.52)	(1.10)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	62.36	38.98	39.01
Observations	11,030	10,925	9,798
Adjusted R^2	0.73	0.54	0.63
Adjusted within R^2	0.04	0.05	0.06
Average Persistence Water or Parks	40%	25%	62%
Average Persistence No Water nor Parks	82%	50%	96%

Notes: All columns contain border-pair and year fixed effects, so all coefficients are estimated on the basis of within placebo D-C pair. The *Post*^{1977} period is 1980-2015. All columns include border-pair and year fixed effects. Water and park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 10% of the area. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without amenities. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.24: Natural amenities (30% threshold) mitigate the persistence of redlining, all D-C neighborhoods within the same MSA

	(1)	(2)	(3)
Dependent variables	% white	% housing units above MSA median home value	% families above MSA median family income
D-graded	-13.14***	-16.20***	-10.70***
	(1.10)	(1.33)	(0.76)
D-graded × Post^{1977}	3.88***	8.93***	2.68***
	(1.41)	(1.78)	(0.94)
Water or park amenities	1.67***	3.66***	1.18**
	(0.61)	(0.82)	(0.54)
Water or park amenities × Post^{1977}	4.00***	0.53	2.40**
	(1.01)	(1.45)	(0.96)
D-graded × Water or park amenities	-0.64	-5.13***	-1.52**
	(1.28)	(0.74)	(0.68)
D-graded × Water or park amenities × Post^{1977}	-0.55	4.29***	0.58
	(1.95)	(1.19)	(0.95)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.29
Adjusted within R^2	0.04	0.07	0.08
Average Persistence Water or Parks	76%	38%	73%
Average Persistence No Water nor Parks	70%	45%	75%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. The *Post*^{1977} period is 1980-2015. Water or park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 30% of the area. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without amenities. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.25: Natural amenities (30% threshold) mitigate the persistence of redlining, bordering D-C neighborhoods

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-7.53***	-9.52***	-5.75***
	(0.97)	(1.01)	(0.62)
D-graded × Post^{1977}	2.82*	6.17***	1.61**
	(1.51)	(1.30)	(0.81)
Water or park amenities	-0.81	-0.58	-1.88**
	(1.18)	(0.87)	(0.86)
Water or park amenities × Post^{1977}	3.46**	2.39	3.68***
	(1.60)	(1.59)	(1.23)
D-graded × Water or park amenities	-1.36	-2.44***	-0.65
	(1.40)	(0.86)	(0.76)
D-graded × Water or park amenities × Post^{1977}	1.93	2.21*	-0.02
	(1.96)	(1.31)	(1.03)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	62.36	38.98	39.01
Observations	11,030	10,925	9,798
Adjusted R^2	0.73	0.53	0.63
Adjusted within R^2	0.04	0.05	0.06
Average Persistence Water or Parks	47%	30%	75%
Average Persistence No Water nor Parks	63%	35%	72%

Notes: All columns contain border-pair and year fixed effects, so all coefficients are estimated on the basis of within placebo D-C pair. The $Post^{1977}$ period is 1980-2015. All columns include border-pair and year fixed effects. Water and park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 30% of the area. Average persistence is computed as the ratio of the D-C gap after the passing of the CRA to the gap before for areas with and without amenities. Family income is only available starting with the 1950 Census columns (1) and (2) are estimated for 1940-2015 and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.26: All D-C neighborhoods and amenities-year fixed effects

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-13.06***	-15.63***	-10.14***
	(1.02)	(1.30)	(0.78)
D-graded × Post^{1977}	3.90**	8.33***	2.33**
	(1.48)	(1.86)	(1.01)
D-graded × Water or park amenities	-0.64	-5.06***	-2.15***
	(1.14)	(0.87)	(0.64)
D-graded × Water or park amenities × Post^{1977}	-0.32	4.60***	1.18
	(1.67)	(1.38)	(0.94)
Area FE	MSA	MSA	MSA
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.28
Adjusted within R^2	0.04	0.07	0.08

Notes: All columns contain MSA and amenities-year fixed effects, so all coefficients are estimated on the basis of all D-C pairs. The $Post^{1977}$ period is 1980-2015. Water and park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 20% of the area. Due to data availability, columns (1) and (2) are estimated for the 1940-2015 period and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.27: Bordering D-C neighborhoods persistence and amenities-year fixed effects

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-7.92***	-9.23***	-5.51***
	(0.99)	(1.05)	(0.67)
D-graded × Post^{1977}	2.44	5.55***	1.15
	(1.48)	(1.35)	(0.94)
D-graded × Water or park amenities	-0.58	-2.63***	-1.11
	(1.17)	(0.79)	(0.70)
D-graded × Water or park amenities × Post^{1977}	2.54*	3.12**	1.06
	(1.49)	(1.37)	(1.06)
Area FE	D-C pair	D-C pair	D-C pair
Mean Dep. Var.	62.36	38.98	39.01
Observations	11,030	10,925	9,798
Adjusted R^2	0.73	0.53	0.63
Adjusted within R^2	0.03	0.05	0.06

Notes: All columns contain border-pair and amenities-year fixed effects, so all coefficients are estimated on the basis of within D-C pairs. The *Post*^{1977} period is 1980-2015. Water and park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features or parks cover at least 20% of the area. Due to data availability, columns (1) and (2) are estimated for the 1940-2015 period and column (3) for 1950-2015. Standard errors are clustered by Census-division and decade ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.28: Parks also have a strong effect in housing values

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
Park amenities	1.50***	3.72***	1.13***
	(0.48)	(1.01)	(0.42)
Park amenities × Post^{1977}	2.25***	-0.57	1.65*
	(0.82)	(1.44)	(0.86)
D-graded × Park amenities	-3.66***	-6.60***	-3.09***
	(1.06)	(1.02)	(0.42)
D-graded × Park amenities × Post^{1977}	-1.38	5.45***	0.31
	(1.37)	(1.35)	(0.62)
Area FE	MSA	MSA	MSA
Water controls	YES	YES	YES
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.29
Adjusted within R^2	0.05	0.08	0.08
Average Persistence Modified	-182%	12%	-24%
Average Persistence Unmodified	66%	64%	82%
Average Persistence Parks	83%	36%	77%
Average Persistence No Water nor Parks	67%	44%	73%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. *Post*^{1977} is defined from 1980-2015. All columns control for water (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the water features), modifications and its interactions with being D-graded and *Post*^{1977}. Water amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features cover at least 20% of the area. Modification is an indicator for waterfront redevelopment projects (1 if the neighborhood falls within the 500 meter buffer around the project, 0 otherwise). Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.29: Waterfront modifications drive the effect of water amenities, as they happen

	(1)	(2)	(3)
Dependent variables	% white	% housing units above MSA median home value	% families above MSA median family income income
D-graded	-12.560*** (1.0081)	-14.705*** (1.2973)	-9.928*** (0.7786)
D-graded × Post^{1977}	4.184*** (1.4313)	8.114*** (1.7330)	2.673*** (0.9963)
Water amenities	1.978* (1.0787)	1.219 (1.2099)	1.548* (0.9003)
Water amenities × Post^{1977}	4.870** (1.9180)	1.799 (2.1013)	0.786 (1.2766)
D-graded × Water amenities	4.560*** (1.5681)	-2.432** (1.0656)	0.317 (1.1642)
D-graded × Water amenities × Post^{1977}	-0.984 (2.9618)	-0.838 (2.0709)	-0.056 (1.7225)
Water amenities × Modification × Post^{1977}	-0.280 (5.7739)	4.845 (4.1092)	0.982 (4.9905)
D-graded × Water amenities × Modification × Post^{1977}	12.959*** (3.6281)	8.424** (4.0234)	9.367* (4.8529)
Area FE	MSA	MSA	MSA
Park controls	YES	YES	YES
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.29
Adjusted within R^2	0.05	0.07	0.08
Average Persistence Modified	-102%	8%	-25%
Average Persistence Unmodified	60%	58%	73%
Average Persistence No Water nor Parks	67%	45%	73%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. *Post*^{1977} is defined from 1980-2015. All columns control for parks (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the parks) and its interactions with being D-graded and *Post*^{1977}. Water amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features cover at least 20% of the area. Modification is an indicator for waterfront redevelopment projects (1 if the neighborhood falls within the 500 meter buffer around the project, 0 otherwise). Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.30: Waterfront modifications drive the effect of water amenities, as they happen

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
Park amenities	1.55***	3.71***	1.09**
	(0.49)	(1.01)	(0.44)
Park amenities × Post^{1977}	2.21***	-0.56	1.65*
	(0.83)	(1.44)	(0.88)
D-graded × Park amenities	-3.58***	-6.69***	-3.15***
	(1.05)	(1.03)	(0.42)
D-graded × Park amenities × Post^{1977}	-1.41	5.60***	0.41
	(1.36)	(1.35)	(0.62)
Area FE	MSA	MSA	MSA
Water controls	YES	YES	YES
Mean Dep. Var.	66.36	44.05	42.82
Observations	22,401	22,172	19,885
Adjusted R^2	0.38	0.24	0.29
Adjusted within R^2	0.05	0.07	0.08
Average Persistence Modified	-102%	8%	-25%
Average Persistence Unmodified	60%	58%	73%
Average Persistence Parks	83%	36%	76%
Average Persistence No Water nor Parks	67%	45%	73%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. *Post*^{1977} is defined from 1980-2015. All columns control for water (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the water features), modifications and its interactions with being D-graded and *Post*^{1977}. Park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around park features cover at least 20% of the area. Modification is an indicator for waterfront redevelopment projects (1 if the neighborhood falls within the 500 meter buffer around the project, 0 otherwise). Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.31: Reduced form results

	(1)	(2)	(3)
Dependent variables	% white	% housing units above MSA median home value	% families above MSA median family income
D-graded	-5.660**	-0.599	-5.542***
	(2.4838)	(1.8869)	(1.5374)
Δ Plague Exposure	88.469	16.474	-2.667
	(112.2846)	(53.2663)	(30.8141)
D-graded $\times$ Δ Plague Exposure	1793.030***	493.545	956.773**
	(683.4471)	(468.8745)	(459.0282)
Δ 1 SD Δ Plague Exposure	19.91	5.48	10.62
Mean Dep. Var.	43.43	38.42	35.68
Observations	1,450	1,450	1,450
Adjusted R^2	0.08	0.39	0.19
Adjusted within R^2	0.03	0.01	0.04

Notes: This table shows the results from regressing the dependent variables in 2015 for the entire D-C sample on a dummy for being D-graded, the experimented change in plague exposure and their interaction. All columns include MSA fixed effects. All columns control for the presence of water, park amenities, waterfront modifications and their respective interactions with being D-graded. Changes in tree canopy are defined as the increase in tree detected pixels during the two periods with aerial imagery. MSA without plagues are excluded. Standard errors are robust and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.32: Second stage

Dependent variables	(1) % white	(2) % housing units above MSA median home value	(3) % families above MSA median family income
D-graded	-71.166***	-18.403	-39.060**
	(24.1802)	(16.7142)	(16.4506)
$\widehat{\Delta TC}$	0.414	0.077	-0.012
	(0.5248)	(0.2490)	(0.1440)
D-graded × $\widehat{\Delta TC}$	27.727***	7.551	14.282**
	(10.1096)	(6.9757)	(6.8543)
Area FE	MSA	MSA	MSA
Amenities and modifications	YES	YES	YES
Mean Dep. Var.	43.43	38.42	35.68
Observations	1,450	1,450	1,450

Notes: This table shows the results from regressing the dependent variables in 2015 on the entire D-C sample on a dummy for being D-graded, the predicted increase in tree canopy and the interaction. The predicted increase in tree canopy is obtained by regressing the increase in tree pixels on a dummy for being D-graded, the experimented plague exposure and the interactions, controlling for the presence of water and park amenities, waterfront modifications and the interactions with redlining. The first stage of this table is in Panel A of Table 6. All columns control for the presence of water, park amenities, waterfront modifications and their respective interactions with being D-graded. Changes in tree canopy are defined as the increase in tree detected pixels during the two periods with aerial imagery. MSA without plagues are excluded. Standard errors are robust and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.33: The effect of waterfront modifications in ownership suggests gentrification

Dependent variables	(1) % black owners	(2) % wite owners	(3) % black families above black MSA median family income
D-graded	8.807***	-10.106***	-8.606***
	(0.9972)	(1.0050)	(1.2680)
D-graded × $Post^{1977}$	-0.318	2.134	0.333
	(1.5781)	(1.4864)	(1.4431)
Water amenities	-2.316**	1.832**	0.381
	(0.9458)	(0.9133)	(1.8954)
Water amenities × $Post^{1977}$	-3.687*	6.305***	1.398
	(2.0825)	(1.9358)	(2.1003)
D-graded × Water amenities	-3.685**	3.583**	1.878
	(1.4306)	(1.4266)	(1.8239)
D-graded × Water amenities × $Post^{1977}$	1.069	-1.467	-0.575
	(2.8888)	(3.0055)	(2.0801)
Water amenities × Modification	-1.428	2.311	-2.515
	(2.4364)	(2.5563)	(3.5137)
Water amenities × Modification × $Post^{1977}$	-0.928	-4.888	-4.729
	(4.4306)	(5.4675)	(4.5819)
D-graded × Water amenities × Modification	-3.688*	3.929*	-5.925
	(1.9606)	(2.0279)	(6.6798)
D-graded × Water amenities × Modification × $Post^{1977}$	-8.646***	11.740***	9.075
	(2.5329)	(3.4447)	(7.6124)
Area FE	MSA	MSA	MSA
Park Controls	YES	YES	YES
Mean Dep. Var.	22.39	71.78	53.92
Observations	22,379	22,379	16,259
Adjusted R^2	0.24	0.32	0.15
Adjusted within R^2	0.04	0.04	0.04
Average Persistence Modified	-451%	-378%	30%
Average Persistence Unmodified	115%	90%	104%
Average Persistence No Water nor Parks	96%	79%	96%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. $Post^{1977}$ is defined from 1980-2015. Water amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features covers at least 20% of the area. Modification is an indicator for waterfront redevelopment projects (1 if the neighborhood falls within the 500 meter buffer around the project, 0 otherwise). All columns control for parks (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the parks) and its interactions with being D-graded and $Post^{1977}$. Ownership shares are computed with respect to occupied housing units for the period 1940-2015. Column (3) is the share of black families with family income above the MSA median black family income, and the estimating period in 1960-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.34: Parks reduce black ownership

Dependent variables	(1) % black owners	(2) % wite owners	(3) % black families above black MSA median family income
Park amenities	-1.18**	1.19**	1.04
	(0.53)	(0.50)	(1.24)
Park amenities × $Post^{1977}$	-1.43*	2.73***	1.04
	(0.76)	(0.81)	(1.46)
D-graded × Park amenities	5.04***	-3.99***	-2.33
	(0.97)	(1.05)	(1.59)
D-graded × Park amenities × $Post^{1977}$	-3.40**	-0.24	-0.37
	(1.62)	(1.55)	(1.77)
Area FE	MSA	MSA	MSA
Water controls	YES	YES	YES
Mean Dep. Var.	22.39	71.78	53.92
Observations	22,379	22,379	16,259
Adjusted R^2	0.24	0.32	0.15
Adjusted within R^2	0.04	0.04	0.04
Average Persistence Modified	-451%	-378%	30%
Average Persistence Unmodified	115%	90%	104%
Average Persistence Parks	73%	87%	100%
Average Persistence No Water nor Parks	96%	79%	96%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. $Post^{1977}$ is defined from 1980-2015. Park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around parks features covers at least 20% of the area. All columns control for water (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the water features), modifications and its interactions with being D-graded and $Post^{1977}$. Ownership shares are computed with respect to occupied housing units for the period 1940-2015. Column (3) is the share of black families with family income above the MSA median black family income, and the estimating period in 1960-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.35: The effect of waterfront modifications in renters suggests gentrification

	(1)	(2)
	% black renters	% white renters
Variables		
D-graded	5.226**	-6.708***
	(2.6029)	(2.5364)
D-graded × Post^{1977}	4.074	-1.847
	(3.0607)	(2.9428)
Water amenities	-1.813	1.497
	(1.0943)	(1.0210)
Water amenities × Post^{1977}	-3.199	5.862***
	(2.0148)	(1.7920)
D-graded × Water amenities	-1.600	1.502
	(1.6160)	(1.6079)
D-graded × Water amenities × Post^{1977}	-1.603	1.106
	(2.8033)	(2.9297)
Water amenities × Modification	-0.333	1.334
	(2.4864)	(2.5652)
Water amenities × Modification × Post^{1977}	-4.339	-3.001
	(4.4242)	(5.2732)
D-graded × Water amenities × Modification	-1.517	2.015
	(2.0740)	(2.0911)
D-graded × Water amenities × Modification × Post^{1977}	-9.881***	12.762***
	(2.6103)	(3.3463)
Area FE	MSA	MSA
Park Controls	YES	YES
Mean Dep. Var.	36.05	56.99
Observations	22,387	22,387
Adjusted R^2	0.44	0.47
Adjusted within R^2	0.03	0.04
Average Persistence Modified	-251%	-277%
Average Persistence Unmodified	168%	114%
Average Persistence No Water nor Parks	178%	128%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. *Post*^{1977} is defined from 1980-2015. Water amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features covers at least 20% of the area. *Post*^{1977} is defined from 1980-2015. Modification is an indicator for waterfront redevelopment projects (1 if the neighborhood falls within the 500 meter buffer around the project, 0 otherwise). All columns control for parks (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the parks) and its interactions with being D-graded and *Post*^{1977}. Renters shares are computed with respect to occupied housing units. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.36: The racial composition of renters does not change with parks

	(1) % black renters	(2) % white renters
Variables		
Park amenities	-1.34* (0.69)	1.37** (0.66)
Park amenities × Post^{1977}	-1.62 (0.98)	2.55** (1.02)
D-graded × Park amenities	4.53*** (1.37)	-3.31** (1.37)
D-graded × Park amenities × Post^{1977}	-2.53 (1.79)	-1.26 (1.66)
Area FE	MSA	MSA
Water controls	YES	YES
Mean Dep. Var.	36.05	56.99
Observations	22,387	22,387
Adjusted R^2	0.44	0.47
Adjusted within R^2	0.03	0.04
Average Persistence Modified	-251%	-277%
Average Persistence Unmodified	168%	114%
Average Persistence Parks	116%	131%
Average Persistence No Water nor Parks	178%	128%

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. $Post^{1977}$ is defined from 1980-2015. Park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around parks features covers at least 20% of the area. All columns control for water (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the water features), modifications and its interactions with being D-graded and $Post^{1977}$. Ownership shares are computed with respect to occupied housing units for the period 1940-2015. Column (3) is the share of black families with family income above the MSA median black family income, and the estimating period in 1960-2015. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.37: The effects of waterfront modifications on share of college graduates suggest gentrification

Variables	(1) % black some college	(2) % white some college
Water amenities × Post^{1977}	2.773***	4.187***
	(0.9613)	(1.1759)
Water amenities × Modification × Post^{1977}	-0.690	-4.248
	(3.2162)	(2.9111)
D-graded × Post^{1977}	-5.342***	-2.882***
	(0.4032)	(0.8010)
D-graded × Water amenities × Post^{1977}	-0.931	-4.722***
	(0.7451)	(1.0301)
D-graded × Water amenities × Modification × Post^{1977}	4.217	14.833***
	(3.8071)	(4.4803)
Area FE	MSA	MSA
Park Controls	YES	YES
Mean Dep. Var.	40.29	52.11
Observations	12084.00	9959.00
Adjusted R^2	0.36	0.28
Adjusted within R^2	0.04	0.03

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. Water amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around water features covers at least 20% of the area. Variables are labeled $Post^{1977}$ to indicate the estimating period is 1980-2015 due to data availability. All columns control for parks (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the parks) and its interactions with being D-graded. Dependent variables are the share of population, black or white, with some college education relative to all population, black or white, aged 25 or older. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.38: Areas with parks have less educated residents

Variables	(1) % black some college	(2) % white some college
Park amenities × Post^{1977}	2.304*** (0.3477)	4.710*** (0.3450)
D-graded × Park amenities × Post^{1977}	-2.661*** (0.4547)	-1.618** (0.6343)
Area FE	MSA	MSA
Water Controls	YES	YES
Mean Dep. Var.	40.29	52.11
Observations	12084.00	9959.00
Adjusted R^2	0.36	0.28
Adjusted within R^2	0.04	0.03

Notes: All columns contain MSA and year fixed effects, so coefficients are estimated on the basis of all D-C neighborhoods within MSA. Park amenities is a dummy variable that takes value one for those neighborhoods in which the 500m buffers around parks features covers at least 20% of the area. All columns control for water (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the water features), modifications and its interactions with being D-graded. Variables are labeled $Post^{1977}$ to indicate the estimating period is 1980-2015 due to data availability. All columns control for parks (a dummy with value one when at least 20% of the neighborhoods' area is covered by the 500m buffer around the parks) and its interactions with being D-graded. Dependent variables are the share of population, black or white, with some college education relative to all population, black or white, aged 25 or older. Standard errors are clustered by Census division-year and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.39: Ownership and tree canopy

Dependent variables	(1) % black owners	(2) % white owners	(3) % black renters	(4) White renters	(5) % black families above black MSA median family income
D-graded	67.686***	-87.445***	73.892***	-80.964***	-55.850***
	(15.1588)	(32.5380)	(13.8394)	(27.8195)	(20.5366)
$\widehat{\Delta TC}$	-0.341	0.621	-0.363	0.522	-0.057
	(0.5979)	(0.6687)	(0.5071)	(0.5669)	(0.1403)
D-graded × $\widehat{\Delta TC}$	-27.418***	35.711**	-29.829***	32.813***	20.630**
	(6.4121)	(14.2240)	(5.8832)	(12.1851)	(9.0193)
Mean Dep. Var.	36.49	51.36	44.52	41.25	49.57
Observations	1,450	1,450	1,450	1,450	1,430

Notes: This table shows the results from regressing the dependent variables on a dummy for being D-graded, predicted tree canopy and the interaction. Predicted tree canopy is obtained by regressing the increase in tree pixels on a dummy for being D-graded, the change in plague exposure and the interaction. The first stage results can be seen in Table 6. Owner and renters shares are computed with respect to occupied housing units for 2015. Column (4) is the share of black families with family income above the MSA median black family income. All specifications include MSA fixed effects. Standard errors are robust and ***, **, * indicate significance at the 1, 5, and 10 percent.

Table 8.3.40: Education and tree canopy

Dependent variables	(1) % black some college	(2) % white some college
D-graded	-9.174	-21.400
	(16.0140)	(19.7313)
$\widehat{\Delta TC}$	0.032	0.035
	(0.1633)	(0.1383)
D-graded × $\widehat{\Delta TC}$	2.477	8.949
	(7.0359)	(8.5988)
Mean Dep. Var.	51.24	49.57
Observations	1,447	1,447

Notes: This table shows the results from regressing the dependent variables on a dummy for being D-graded, predicted tree canopy and the interaction. Predicted tree canopy is obtained by regressing the increase in tree pixels on a dummy for being D-graded, the change in plague exposure and the interaction. The first stage results can be seen in Table 6. Dependent variables are the share of population, black or white, with some college education relative to all population, black or white, aged 25 or older. All specifications include MSA fixed effects. Standard errors are robust and ***, **, * indicate significance at the 1, 5, and 10 percent.