

ARTICLE OPEN



Residential housing segregation and urban tree canopy in 37 US Cities

Dexter H. Locke¹✉, Billy Hall², J. Morgan Grove¹, Steward T. A. Pickett³, Laura A. Ogden⁴, Carissa Aoki⁵, Christopher G. Boone⁶ and Jarlath P. M. O'Neil-Dunne⁷

Redlining was a racially discriminatory housing policy established by the federal government's Home Owners' Loan Corporation (HOLC) during the 1930s. For decades, redlining limited access to homeownership and wealth creation among racial minorities, contributing to a host of adverse social outcomes, including high unemployment, poverty, and residential vacancy, that persist today. While the multigenerational socioeconomic impacts of redlining are increasingly understood, the impacts on urban environments and ecosystems remain unclear. To begin to address this gap, we investigated how the HOLC policy administered 80 years ago may relate to present-day tree canopy at the neighborhood level. Urban trees provide many ecosystem services, mitigate the urban heat island effect, and may improve quality of life in cities. In our prior research in Baltimore, MD, we discovered that redlining policy influenced the location and allocation of trees and parks. Our analysis of 37 metropolitan areas here shows that areas formerly graded D, which were mostly inhabited by racial and ethnic minorities, have on average ~23% tree canopy cover today. Areas formerly graded A, characterized by U.S.-born white populations living in newer housing stock, had nearly twice as much tree canopy (~43%). Results are consistent across small and large metropolitan regions. The ranking system used by Home Owners' Loan Corporation to assess loan risk in the 1930s parallels the rank order of average percent tree canopy cover today.

npj Urban Sustainability (2021)1:15; <https://doi.org/10.1038/s42949-021-00022-0>

INTRODUCTION

Spatial, social, and environmental inequities pose significant challenges for American cities^{1,2}. Urban inequity is the result of historical and systemic forces, including structural racism and segregation, which have enduring effects on the ways cities function socially, economically, and ecologically^{3–5}. For instance, decades of racial discrimination in housing policy created barriers to homeownership, employment, and access to quality education for people of color, making it difficult to build wealth across generations^{6–8}. While the mechanisms linking structural racism to wealth creation and socioeconomic status are well understood⁹, it is less clear how housing segregation may have played a role in shaping urban ecosystems.

This paper investigates how the historic practice of redlining, one of the most consistent, wide-spread, spatial, and racial forms of US housing practices, relates to the contemporary distribution of urban tree canopy, commonly understood as a vital component of urban ecosystem health and sustainability^{10,11}. Trees provide a host of ecosystem services and social benefits, including heat island mitigation^{12,13}. In the United States, ~1500 heat-related deaths occur each year¹⁴, and the impact of heat stress is likely to increase given current climate projections¹⁵. Existing tree canopy cover¹³ and the replacement of impervious surfaces with tree canopy can lower urban temperatures¹² and save lives. But trees and tree canopy are not distributed equitably^{16–18}. Recent meta-analyses show that lower-income urban areas¹⁶ and areas with more racial minorities¹⁷ have less tree canopy cover, an environmental injustice that can exacerbate health problems for already disadvantaged groups.

These racial and geographic disparities in urban tree canopy parallel other striking patterns of racialized environmental inequity documented by environmental justice (EJ) research. For more than three decades, EJ researchers have developed an enormous body of evidence highlighting the disproportionate concentration of environmental hazards and burdens in communities of color, and conversely, the privileged access to environmental amenities in predominantly white communities^{19–22}. More recent waves of this scholarship have begun to explicitly connect racial disparities in environmental “goods” and “bads” to structural racism and discrimination in policy^{23,24}. Institutionalized policies and practices intended to racially segregate (and concentrate wealth in white communities) have been increasingly shown to produce racially uneven landscapes of environmental privilege and risk, even decades later^{2,25–27}.

In thinking about urban tree canopy, the space to plant trees is often a legacy of the urban built environment, which in the United States stems from histories of deliberate and systematic racial discrimination in housing and urban development^{26,28}. In 1933, the US Congress created the Home Owners' Loan Corporation (HOLC) to assist Americans struggling to pay their mortgages in the wake of the Great Depression. To guide lending criteria, the HOLC developed neighborhood appraisal maps for 239 urban areas, ranking the perceived risk of investing in particular neighborhoods using a color-coded scale of “A” (green), “B” (blue), “C” (yellow), and “D” (red)²⁹. Appraisals were based primarily on an area's demographic characteristics and the age and physical condition of its housing stock. Areas with predominantly U.S.-born, white populations, and newer housing stock were often codified as the “safest” places for banks to

¹USDA Forest Service, Northern Research Station, Baltimore Field Station, Baltimore, MD, USA. ²National Socio-Environmental Synthesis Center (SESYNC), Annapolis, MD, USA. ³Cary Institute of Ecosystem Studies Box AB, Millbrook, NY, USA. ⁴Department of Anthropology, Dartmouth College, Hanover, NH, USA. ⁵Bates College, Environmental Studies Program, Lewiston, ME, USA. ⁶School of Sustainability, Arizona State University, Tempe, AZ, USA. ⁷Spatial Analysis Laboratory, Rubenstein School of Environment & Natural Resources, University of Vermont, Burlington, VT, USA. ✉email: dexter.locke@gmail.com

invest and were graded “A” and “B.” Meanwhile, areas with somewhat older structures and/or a presence of foreign-born residents were commonly ascribed a “C” grade, while areas with significant numbers of racial and ethnic minorities, foreign-born residents, families on relief, and having older housing were almost always viewed as “hazardous” and given the lowest grade, “D.” The term “redlining” is used because areas graded “D” were shaded red on the HOLC maps. In effect, while race was not the only criterion considered in designating grades, the maps formally embedded race into neighborhood appraisal processes by systematically factoring in the race of an area’s occupants into the perceived long-term value of an area^{30,31}.

Some context is important to better understand HOLC’s residential security maps and its practices from 1934 to 1951. While the HOLC created uniform guidelines for neighborhood appraisal, because appraisals were produced in direct consultation with local municipal officials, loan officers, appraisers, and realtors, evidence suggests some variation in the grading across cities³². Still, these agents were familiar with their city’s specific patterns of residential segregation. More importantly, many local actors were already part of the power structures that had created, maintained, or profited from the prevailing racist housing policies and practices. These policies and practices included segregation ordinances, racially-restrictive deed covenants, and zoning plans that promoted their agendas of racial and immigrant exclusion^{33–37}. Thus, the HOLC maps helped codify the local real estate industry’s consensus of perceived neighborhood value, which often institutionalized existing local inequities in borrowers’ access to credit^{31,38}.

The extent to which the HOLC’s maps, guidelines, and practices influenced the actual distribution of mortgages remains uncertain. Some evidence suggests that lending practices varied by lender and geography, despite the HOLC’s systematic guidelines³². In addition, some have argued that it is unlikely that the Federal Housing Administration (FHA), which issued long-term mortgages, cooperated directly with the HOLC³². Yet it has been demonstrated that the FHA overwhelmingly prioritized granting mortgages for new homes, which would have been located in areas graded “A” by the HOLC. For instance, between 1934 and 1962, the FHA and the Veterans Administration lent over \$120 billion for new housing, and 98% of this money was distributed to white residents compared to <2% for African Americans and other people of color^{39,40}. During this period, African Americans represented ~10% of the US population⁴¹. Studies in Houston and Boston show that even when controlling for income, whites were nearly three times as likely to receive a mortgage loan^{3,42}.

A large consensus among housing policy scholars is that the federal government helped institutionalize a two-tiered, racialized lending system. One tier provided federally-backed mortgages to higher-graded neighborhoods with predominantly U.S.-born, affluent, white populations occupying newer housing stock. A second-tier subjected residents in the “yellow” and “red” neighborhoods, which housed predominantly low-income African Americans and immigrants in older buildings. Homeowners in the second tier experienced predatory lending schemes or no mortgage lending at all^{35,43}. For decades, many whites benefited from privileged access to credit, home ownership, and wealth accumulation based on home equity, while African Americans were largely denied this route to economic prosperity^{3,33,44–46}. Redlining created systematic disinvestment in minority communities that were located in the denser, older urban core while protecting the property values and resources of white communities moving into desirable homes in the suburbs. Indeed, the post-World War II suburban development supported by federal subsidies created new, exclusively white geographies that generated enormous new wealth^{31,47}.

Although Congress officially outlawed racial discrimination in housing with the Fair Housing Act of 1968, studies continue to document its enduring effects. Many formerly redlined areas continue to struggle with segregation, poverty, unemployment, low educational attainment, and poor health outcomes today^{3,48,49}. Research shows that compared to areas receiving higher grades by the HOLC, lower-graded areas exhibit declines in home ownership, housing value, and credit scores⁵⁰.

Despite the abundance of evidence on the social and economic impacts of racist housing policy, little is known about the relationships among redlining, social disadvantage, and environmental quality. It is well-documented that various social disadvantages are bundled in racially segregated urban areas (3), and environmental justice scholars have demonstrated that these outcomes are the product of profound historical and present-day racist and discriminatory policies and institutions, such as redlining, blockbusting, and zoning^{51–55}. Environmental justice scholars and activists have convincingly argued that fair processes governed by just institutions are equally if not more important than equitable environmental outcomes because process change can lead to enduring systemic change^{56,57}.

It has also been documented that lower-income areas¹⁶ and areas with more racial minorities¹⁷ tend to have less tree canopy cover. However, the relationships among long-term discriminatory housing practices and contemporary environmental conditions remain poorly understood. The distribution of current urban tree canopy cover offers one perspective on environmental inequities related to housing segregation.

Research in Baltimore, MD has shown that redlining and other racially-biased housing practices have historically shaped the location of investments in environmental amenities such as trees and parks and the allocation of environmental disamenities via non-conforming zoning^{2,38,58,59}. Redlined, African American neighborhoods of East and West Baltimore, graded D in the HOLC system, had overcrowded and poor quality housing and higher exposure to noise and other pollution from nearby industries². These denser, D-graded areas had less available space for trees and tree planting, while A-graded areas comprised of single-family homes on larger lots could maintain, grow, and plant additional trees. Race-based evaluations of credit-worthiness also shaped access to wealth accumulation and related political power. Residents in A-graded areas directed municipal investments into street tree plantings, creating public parks with trees, and invested their own resources into trees on their private lands^{26,59}. At the same time, residents in D-graded areas had less access to public investments and were more likely to spend their lower wages on other necessities such as rent, food, or transportation. Thus, differences in lot sizes, money, and access to power along HOLC neighborhood lines played an important role in shaping the distribution of Baltimore’s urban tree canopy over the long term².

Our goal in this paper is to examine whether there are similar patterns in the distribution of tree canopy by HOLC-graded neighborhoods in other cities. These analyses are possible because redlining was a national process, initiated by the Federal government in collaboration with state and local governments. It was a practice that was spatially-explicit and applied to 239 cities in the same time period throughout the country. These characteristics make the practice of redlining particularly well-suited for within- and cross-city comparisons. We examined whether historic redlining is statistically associated with contemporary spatial distributions of tree canopy for a range of metropolitan areas across a spectrum of area, population, and climate. This paper assesses whether there are differences in current tree canopy cover among historic HOLC classes and whether differences among HOLC classes are consistent among cities.

RESULTS

There is a strong relationship between HOLC grades and tree canopy: areas formerly graded D have 21 percentage points less tree canopy than areas formerly graded A. One-way ANOVA showed significant differences in tree canopy by HOLC grade [$F(3, 3184) = 253.9, p < 0.001$]. Post hoc comparisons using the Tukey HSD test indicated that the same hierarchical ranking system used by HOLC to assess loan risk in the 1930s is paralleled by the rank order of average percent tree canopy cover today. Areas formerly graded D have significantly less tree canopy ($M = 20.9$ percentage points, $SD = 12.2$), than areas graded C ($M = 24.6, SD = 10.9$), B ($M = 32.4, SD = 13.8$), or A ($M = 41.1, SD = 14.7$). All six pairwise combinations were significantly different at the $p < 0.0001$ level. The same model was re-fit as a linear regression so that areas graded A are the reference, with differences in means as estimated coefficients, as a baseline model (Table 1, Model 1).

To test for unobserved city-specific factors, a separate unconditional one-way ANOVA was performed. This second ANOVA showed significant differences in tree canopy by city [$F(36, 3151) = 21.60, p < 0.001$]. The intraclass correlation coefficient (ICC) indicated that 23% of the variance in tree canopy cover was from city to city (Table 1, Model 2). A mixed effects model with fixed effects for HOLC grade and random effects for city (Table 1, Model 3) showed that the areas given less-favorable grades by HOLC have significantly less canopy cover than their higher-graded counterparts, with overlap between C and D areas (i.e., $D \leq C < B < A$). Comparing the three model specifications (fixed effects for HOLC grade only, random effects for city, and a specification with both HOLC fixed effects and city random effects) with an AIC-minimization criterion showed that Model 3's added complexity provided the best fit (Table 1). Model 3's regression-adjusted estimates of tree canopy cover suggest that areas formerly graded D had 21 percentage points less tree canopy ($\gamma_{30} = -20.79, 95\% [-22.27, -19.31]$) (or 22% cover) than areas formerly graded A ($\gamma_{00} = 43.44, 95\% [40.80-46.07]$) and the HOLC categories explained 19% of the tree canopy variance while city-to-city variation explained an additional 25%.

Results of further tests and robustness checks

A-graded neighborhoods were often the rarest, making within-city analyses under-powered statistically. In cities with 10 or more A-graded neighborhoods (Fig. 1), within-city analyses of tree canopy cover by grade confirmed the pooled analyses' findings (Fig. 2). Wilcoxon tests showed lower median tree canopy in D neighborhoods compared to A neighborhoods,

except in Seattle ($p = 0.093$). Although the sample sizes for many cities do not permit statistical analyses of within-city analyses of canopy by HOLC grade, the boxplots in Fig. S1 illustrate variation among classes within each city. Tree canopy today is almost always in rank order of HOLC grades.

It is possible that the main results reported in Model 3 are driven by the patterns and sample size in the largest 16 cities with at least 50 HOLC-defined neighborhoods. We therefore re-fit Model 3 excluding the largest 16 cities and the results were substantively the same (Table S1); formerly D-graded areas have about 23% tree canopy, while formerly A-graded areas have nearly twice as much canopy today (43%). Therefore, the findings are not attributable to the patterns found in the largest cities.

DISCUSSION

The link between redlining and socioeconomic outcomes such as poverty and home foreclosure has previously been documented^{3,29,33,44-46,48-50,60}. The lack of access to wealth via homeownership had a powerful influence on real estate markets. People of color were deprived of an important path to wealth accumulation in many urban areas across the US^{29,33,34}. However, the relationships among historic discriminatory housing practices and current environmental conditions remain poorly understood. Redlining was one of the most consistent, wide-spread, spatial, and racial forms of US housing practices. The relationship between redlining and the current distribution of urban tree canopy cover offers a preliminary window into these larger, long term, and complex dynamics. Our research supports prior work on social disparities corresponding to redlining grades by adding evidence pertaining to environmental inequities.

Trees are an important component of the urban environment. They reduce the urban heat island effect^{12,13} and provide a number of other public health benefits⁶¹ such as crime reduction⁶². In order to consider whether historic social disparities are paralleled by contemporary disparities in tree canopy, this paper examined variations in tree cover by HOLC-defined neighborhoods and the metropolitan regions containing those neighborhoods. The difference was significant: formerly D-graded areas have about 23% tree canopy today while formerly A-graded areas have nearly twice as much (43%). We found that just two variables, HOLC neighborhood grade and city, explained 43% of the variance (Table 1).

To be very clear, this study used a cross-sectional, observational quantification of social-ecological patterns that is fundamentally incapable of finding, identifying, and/or ascribing causality for

Table 1. Regressions, for % tree canopy cover, by Home Owners Loan Corporation and metropolitan region.

	Model 1: Fixed effects of HOLC grade			Model 2: Random effects of city			Model 3: Mixed effects		
	Estimates	95% CI	p	Estimates	95% CI	p	Estimates	95% CI	p
<i>Fixed effects</i>									
(Intercept) HOLC Grade A: "Best" (γ_{00})	41.04	39.64 to 42.45	<0.001	29.31	27.03 to 31.59	<0.001	43.44	40.80 to 46.07	<0.001
HOLC Grade B: "Still Desirable" (γ_{10})	-8.66	-10.31 to -7.02	<0.001				-9.06	-10.51 to -7.61	<0.001
HOLC Grade C: "Definitely Declining" (γ_{20})	-16.41	-17.96 to -14.85	<0.001				-16.83	-18.21 to -15.45	<0.001
HOLC Grade D: "Hazardous" (γ_{30})	-20.11	-21.77 to -18.44	<0.001				-20.79	-22.27 to -19.31	<0.001
<i>Random effects</i>									
σ^2				153.42			116.42		
τ_{00}				46.45 city			50.88 city		
ICC				0.23 city			0.30 city		
Observations	3188			3188			3188		
R^2 /adjusted R^2	0.193/0.192			0.000/0.232			0.187/0.434		
AIC	25,084.874			25,202.044			24,336.329		

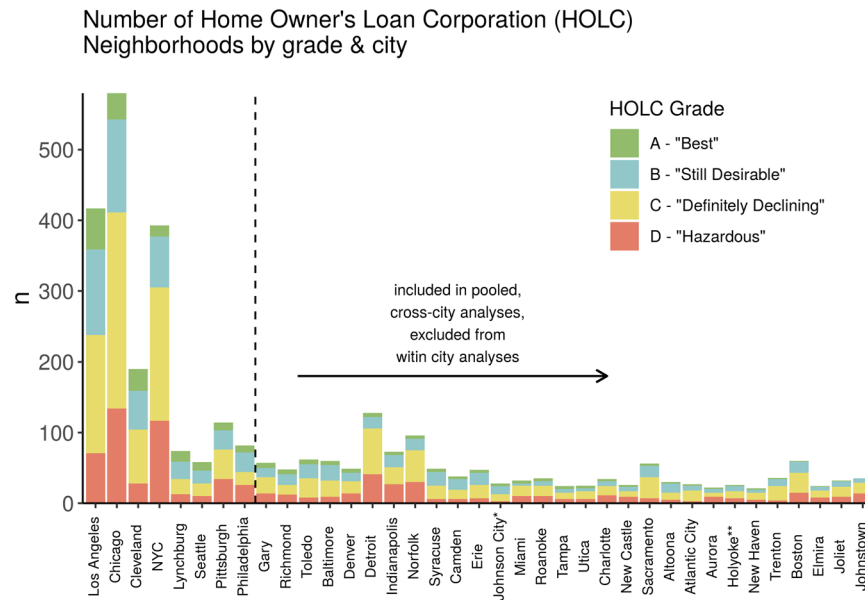


Fig. 1 The distribution of HOLC neighborhoods by type and city shows the overall within-city and across-city patterns. Larger and more segregated metropolitan areas tend to have more HOLC-defined neighborhoods. Cities are sorted by the number of A-graded neighborhoods. Only eight cities have ≥ 10 Grade-A neighborhoods (left) to permit within-city analyses. In the main analysis, all neighborhoods are used. *Johnson City/Birmingham, NY; **Holyoke/Chicopee, MA.

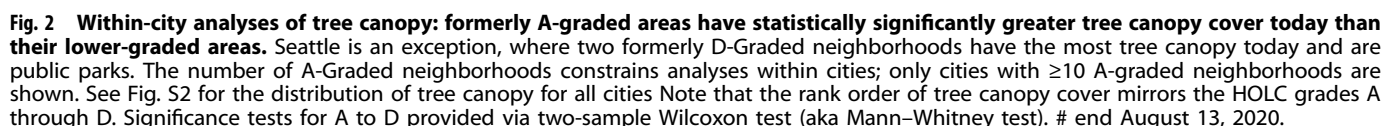
complementary or competing explanations of the process. The findings are consistent with other recent examinations of HOLC grades and vegetation cover⁶³, the urban heat island effect⁶⁴, and even premature births⁶⁵: formerly D-graded areas on average have less vegetation, are hotter, and are associated with statistically significantly more preterm births⁶⁵. The determinants of tree canopy cover in urban areas are complex²⁸. Our paper highlights one possible factor that may have played a role while also ruling out random chance. We argue that redlining is an understudied process in urban ecology and that our findings suggest that the role of redlining in shaping tree canopy, in concert with other explanatory factors, warrants further process-based research. HOLC's redlining was a moment in a long-term history of discriminatory housing practices in the United States³¹. Thus, in-depth and comparative research is needed to understand the systemic processes among long-term discriminatory housing practices and contemporary environmental conditions².

There may be several systemic explanations for our pattern-based results. If redlining reflected existing differences in lot size and reinforced those differences through preferred investment over the long term, we could expect to see more extensive contemporary tree canopy within formerly A-graded areas. This contemporary distribution of canopy cover may be due in part to the fact that residential lots in these areas would have been larger and had more space for trees. A-graded areas were also more affluent, and households may have had higher disposable incomes to invest in landscaping such as trees. Further, because redlining helped shape wealth accumulation and related political power by race and geography, the privilege of those living in formerly A-graded neighborhoods may have served to direct public investments in tree canopy over the long term for street trees and trees in parks or through continued private household investment in landscaping on their own larger residential properties^{66–69}. In this way, complex and reinforcing positive feedback loops may have occurred, perpetuating relationships among housing markets, affluence, race, and trees. Such a positive feedback loop may have also been mirrored in formerly D-graded areas with lower tree canopy today due to smaller lots, industrial

land uses not conducive to tree canopy cover, fewer resources for maintaining trees on properties, and less influence over public investments over the long term. Our results are consistent with both of these rationales. A process-based study is beyond the scope of this paper, but our findings provide a robust starting point to examine the longitudinal dynamics between redlining and tree canopy cover.

Our results point to at least three other areas that could benefit from further research. First, more research may be needed to understand the mechanisms for why the strong association between HOLC categories and urban tree canopy exists some 80 years after the HOLC maps were drawn and the roles that different actors may have played to maintain these differences. Many A-graded areas were suburban areas that had been zoned for single-family housing with large lot sizes⁷⁰. D-graded areas had denser housing stock, but they may have also contained non-residential land uses, such as industrial sites, which might have been unfavorable for trees. A next step could be to examine different residential densities, land uses associated with different jurisdictions, policies, and tree planting programs in the different cities over time. For example, D-graded areas could have been more susceptible to urban renewal projects, supporting highways, and other large-scale infrastructure projects that could have required tree removals or made space for new trees. Analyses of changing land uses, local policies, demographic trends, or historic aerial imagery could enable a greater understanding of the extent to which HOLC grades 'locked in' urban forms that are more or less amenable to tree canopy.

A second approach would be to examine areas that do not match the overall pattern. So-called statistical "deviant case analyses"⁷¹ may help to build better theory about spatial, social, and environmental inequities, including historic processes of urban renewal and contemporary processes of gentrification and climatic conditions. For example, tree canopy cover in Seattle, WA in formerly A-graded neighborhoods is generally greater than in formerly D-graded neighborhoods (Fig. 2), but the differences were not statistically significant ($p = 0.093$). Moreover, the two areas with the highest percent of tree canopy cover in Seattle



Given the long history of disinvestment in African-American communities in the United States, we sought to understand the extent to which a program in the 1930s that altered the distribution and flow of land and capital along racial lines is associated with contemporary tree canopy cover in urban areas.

Land cover data for 23 areas were downloaded from The Spatial Analysis Lab (The SAL, <http://gis.w3.uvm.edu/utc/>, Table S2) at the University of Vermont. The SAL routinely maps large spatial extents such as counties

and their methods are detailed elsewhere^{82–84}. Next, tree canopy data for the entire state of Pennsylvania were obtained for all HOLC-mapped cities in Pennsylvania from SAL (Altoona, Johnstown, New Castle, Philadelphia, and Pittsburgh, <http://letters-sal.blogspot.com/2015/09/pennsylvania-statewide-high-resolution.html>). Tree canopy data for eight cities (Baltimore, MD; Johnson City-Binghamton, Syracuse, and Utica, NY; Lynchburg, Norfolk, Richmond, and Roanoke, VA) were obtained (Chesapeake Bay Program, <https://chesapeakeconservancy.org/conservation-innovation-center/high-resolution-data/>). Data for New Jersey (Atlantic City, Camden, and Trenton) were obtained (Pennsylvania Spatial Data Access, <http://www.pasda.psu.edu/uci/DataSummary.aspx?dataset=3193>). Finally, a literature review was used to identify ($n=8$) sources for additional land cover data overlapping HOLC-graded areas and corresponding authors were contacted for data access (Los Angeles and Sacramento, CA; Denver, CO; Miami and Tampa, FL; Holyoke-Chicopee, MA; Toledo, OH; and Seattle, WA). In total, there were 3188 HOLC-defined neighborhoods, from 37 cities, in 16 states from 11 sources (Table S2). Statistical analyses were conducted in R v. 3.6.1⁸⁵ using the tidyverse⁸⁶, simple features⁸⁷, ggpubr⁸⁸, lme4⁸⁹, sjPlot⁹⁰, and sjstats⁹¹ packages.

Dependent variables

The dependent variable was the percentage of tree canopy cover within each HOLC zone. Consistent with previously published literature^{18,92}, we define and operationalize tree canopy as “the layer of leaves, branches, and stems of trees that cover the ground when viewed from above”⁹³. After projecting the HOLC polygons obtained from the Mapping Inequality Project to match the land cover data, the Tabulate Area tool was used in ArcMap Version 10.2.2 (ESRI, 2014) to calculate the percent of tree canopy cover for each polygon. In seven cities (Boston, Denver, Detroit, New Haven, New York City, Seattle, and Toledo), tree canopy data were not available for the entire extent of the HOLC-defined neighborhoods, which occasionally extended into suburban areas surrounding the municipalities of interest and 156 polygons had to be omitted. This represents 4.67% of the dataset and was unavoidable. As a robustness check, described below, our main regression model was re-fit with those seven cities entirely removed.

Empirical strategy

We conducted two analyses of variance (ANOVA) with tree canopy as the dependent variable. In the first ANOVA, the independent variable was the HOLC categories in order to test our main hypothesis that mean canopy cover varied by grade. A post hoc Tukey HSD was then used to examine which pairs of grades differed from each other. This initial ANOVA was re-fit as a linear regression model so that Grade A would be the base-case for comparison, and letters B, C, and D would be estimated as differences in means from A. This is Model 1.

In the second ANOVA, the independent variable was the city in which each neighborhood was located (hereafter Model 2). This analysis was conducted because we were concerned that unobserved city-specific characteristics pertaining to such things as land use policy, urban form, climate, and other factors may have influenced tree canopy cover. The purpose of Model 2 was to test whether tree canopy cover varied across each study city.

As anticipated, tree canopy varies significantly by city. We therefore fit a mixed-effect model with the four-category HOLC grades as the fixed effects, with random intercepts for city, as shown in Eq. (1) and termed Model 3.

$$\eta_{ij} = \gamma_{00} + \gamma_{10}\text{HOLC}_{\text{grade}_B} + \gamma_{20}\text{HOLC}_{\text{grade}_C} + \gamma_{30}\text{HOLC}_{\text{grade}_D} + \mu_{0j} + e_{ij} \quad (1)$$

Where η_{ij} is tree canopy as a percentage land area for HOLC polygon i in city j . HOLC grade A is the reference, and γ_{00} is the intercept and mean value of percent tree canopy cover in formerly A-graded neighborhoods. γ_{10} , γ_{20} , γ_{30} , are the coefficients of interest, which represent the differences in mean tree canopy from A by HOLC grades B, C, and D, respectively. μ_{0j} represents the city-specific random intercept, which was included to capture unobserved aspects of each city, e_{ij} is the observation-level residuals, σ^2 is the within-city variance, and τ_{00} represents the variance across cities. The variance partitioning coefficient, also known as the intraclass correlation coefficient (ICC) is “a population estimate of the variance explained by the grouping structure”⁹⁴, which was calculated as the between-group-variance (τ_{00} , random intercept variance) divided by the total variance (i.e., sum of between-group-variance τ_{00} and within-

group σ^2 residual variance), shown in Eq. (2).

$$\text{ICC} = \tau_{00} / [\tau_{00} + \sigma^2] \quad (2)$$

T-statistics were treated as Wald Z-statistics for calculating the confidence intervals and p -values, assuming a normal-distribution. An approximate R^2 was computed as the proportion of variance explained in the random effect after adding the categorical HOLC fixed effect to the model. This is computed as the correlation between fitted and observed values⁹⁵. AIC minimization was used to compare Models 1, 2, and 3, and to determine the best fitting model⁹⁶.

Cities with enough A- and D-graded neighborhoods were examined in order to determine if the patterns from cross-city, pooled analyses hold within individual cities. D-graded areas are common, but A-graded areas were limiting. For each city with ≥ 10 HOLC-defined A-neighborhoods ($n=8$: Los Angeles, Chicago, Cleveland, New York City, Lynchburg, Seattle, Pittsburgh, Philadelphia), Wilcoxon rank-sum tests were used to compare pairwise differences in tree canopy cover from A to D neighborhoods. All other pairwise tests were omitted for parsimony (Fig. 2).

Methods for further tests and robustness checks

Four types of checks were conducted: one set to assess the potentially undue influence of cities with many HOLC-defined neighborhoods, a second to assess the influence of metropolitan areas with partially missing data, and a third to examine the sensitivity of grouping the five boroughs of New York City, and Chelsea and Cambridge with Boston, and a fourth to examine data from different sources.

Two strategies were used in order to evaluate whether the results of Models 1, 2, and 3 were driven by the metropolitan areas with the most HOLC-defined neighborhoods. First, the boxplots for all cities are provided in Fig. S1 so that the within-city patterns can be examined visually. Second, as a robustness check, Model 3 was re-fit without data from the metropolitan areas with ≥ 50 neighborhoods to see if the patterns would still hold (Table S1). The inferences from this smaller model remain unchanged, however, the confidence intervals are larger by construction.

Tree canopy data were not available for the entire extent of the HOLC-defined areas in seven metropolitan areas. The missing data are usually at the edges of the geographic extent, and therefore non-random. Specifically, tree canopy data were not available for the entire extent HOLC-defined neighborhoods in Boston, Denver, Detroit, New Haven, New York City, Seattle, and Toledo, which collectively represent 4.67% of the total dataset's observations. To address non-random, partially missing data at the edges of these metropolitan regions, Model 3 was re-fit with these cities removed entirely (Table S1, Model 5). Model 5 provides substantively similar results and interpretation to the main Model 3 and the point estimates remain within the bounds of Model 3's confidence intervals.

The sensitivity of the analytical decision to group the five boroughs of New York City, and Chelsea and Cambridge with Boston was also examined. A version of Model 3 (Table S1, Model 5) was fit without grouping, which adds 6 additional random intercepts. Again, no substantive changes were observed.

Finally, land cover data for Sacramento, Denver, Miami, Tampa, Holyoke-Chicopee, Toledo, and Seattle all came from different sources (Table S1, Model 6). It is possible that data from those cities may have influenced the results if the land cover data were not comparable to those produced by SAL. Based on Model 6, no substantive changes were observed. All robustness check models supported the inferences of the main results: formerly D-graded areas had roughly half as much tree canopy as formerly A-graded areas.

Limitations

Cross-city analyses⁹⁷ and meta-analyses^{16,17} have demonstrated inequitable distribution of tree canopy by already disadvantaged groups. This paper builds on those studies by using a consistent approach across 37 cities. These 37 cities (or $\sim 15\%$ of all redlined cities) were chosen based on availability of data. However, this convenience sample nevertheless covers a range of characteristics in population from $\sim 42,000$ people (Lynchburg, VA) to ~ 7.2 million people (New York City) at the time they were redlined in 15 states. When they were redlined, these 37 urban areas analyzed housed ~ 28.7 million people.

DATA AVAILABILITY

All data generated and analyzed as part of this study are openly available from the Environmental Data Initiative (EDI) Data Portal via the following <https://doi.org/10.6073/pasta/4ccbc7087959dc2a25063e589dee7718>⁹⁸. The data are as follows: (1) City-specific file geodatabases with feature classes of the HOLC polygons obtained from the Mapping Inequality Project <https://dsl.richmond.edu/panorama/redlining/>, and tables summarizing tree canopy, and in some cases other land cover classes. (2) An *.R script that replicates all of the analyses, graphs and tables in the article describing the related study. Other double checks, exploratory and miscellaneous outputs can also be created by the script. (3) A *.csv file containing city, the HOLC grade, and the percent tree canopy cover. This can be used to create the main findings of the article and this flat file is provided as an alternative to running the R script to extract information from the geodatabases, combine and analyze them. The intention is that this file is more widely accessible; the underlying information is the same.

Received: 20 March 2020; Accepted: 8 January 2021;

Published online: 25 March 2021

REFERENCES

- Sampson, R. J. Urban sustainability in an age of enduring inequalities: Advancing theory and econometrics for the 21st-century city. *Proc. Natl. Acad. Sci.* **114**, 1–6 (2016).
- Grove, J. M. et al. The Legacy Effect: Understanding How Segregation and Environmental Injustice Unfold over Time in Baltimore. *Ann. Am. Assoc. Geogr.* **108**, 524–537 (2018).
- Massey, D. & Denton, N. *American Apartheid: Segregation and the Making of the Underclass* (Harvard University Press, 1993).
- Sampson, R. J. & Wilson, W. J. in *Race, Crime, and Justice: A Reader* (eds Hagan, J. & Peterson, R.). 37–54 (Stanford University Press, 2005).
- Frazier, J. *Race and Place: Equity Issues in Urban America* (Routledge, 2003).
- Collins, W. J. & Margo, R. A. Residential segregation and socioeconomic outcomes: When did ghettos go bad? *Econ. Lett.* [https://doi.org/10.1016/S0165-1765\(00\)00300-1](https://doi.org/10.1016/S0165-1765(00)00300-1) (2000).
- Oliver, M. & Shapiro, T. M. *Black Wealth/White Wealth: A New Perspective on Racial Inequality* 2nd edn. <https://doi.org/10.4324/9780203707425> (2013).
- Sharkey, P. *Stuck in Place: Urban Neighborhoods and the End of Progress Toward Racial Equality* (University of Chicago Press, 2013).
- Harvey, D. *Social Justice and the City* (University of Georgia Press, 1973).
- Hewitt, V., Mackres, E. & Shickman, K. Cool policies for cool cities: Best practices for mitigating urban heat islands in North American cities. *ACEEE Summer Study Energy Effic. Build.* **10**, 131–145 (2014).
- City of New York. *PlaNYC: A Greater, Greener New York* (2007).
- Ziter, C. D., Pedersen, E. J., Kucharik, C. J. & Turner, M. G. Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer. *Proc. Natl. Acad. Sci. USA* **116**, 7575–7580 (2019).
- McDonald, R. I., Kroeger, T., Zhang, P. & Hamel, P. The value of US urban tree cover for reducing heat-related health impacts and electricity consumption. *Ecosystems* **23**, 137–150 (2019).
- Epstein, P. R. et al. *Climate Change Futures: Health, Ecological and Economic Dimensions* (2005).
- Luber, G. & McGeehin, M. Climate change and extreme heat events. *Am. J. Prev. Med.* <https://doi.org/10.1016/j.amepre.2008.08.021> (2008).
- Gerrish, E. & Watkins, S. L. The relationship between urban forests and income: a meta-analysis. *Landsc. Urban Plan.* **170**, 293–308 (2018).
- Watkins, S. L. & Gerrish, E. The relationship between urban forests and race: a meta-analysis. *J. Environ. Manage.* **209**, 152–168 (2018).
- Schwarz, K. et al. Trees grow on money: urban tree canopy cover and environmental justice. *PLoS ONE* **10**, e0122051 (2015).
- United Church of Christ. *Toxic Wastes and Race in the United States: A National Report on the Racial and Socio-economic Characteristics of Communities with Hazardous Waste Sites* (1987).
- Bullard, R. D. *Dumping in Dixie: Race, Class, and Environmental Quality* 3rd edn. <https://doi.org/10.4324/9780429495274> (2018).
- Cole, L. W. & Foster, S. R. *From the Ground Up: Environmental Racism and the Rise of the Environmental Justice Movement* (New York University Press, 2000).
- Bullard, R. D., Mohai, P., Saha, R. & Wright, B. Toxic wastes and race at twenty: why race still matters after all of these years. *Environ. Law* **38**, 371 (2008).
- Taylor, D. E. *Toxic Communities: Environmental Racism, Industrial Pollution, and Residential Mobility* (New York University Press, 2014).
- Boone, C. G., Buckley, G. L., Grove, J. M. & Sister, C. Parks and people: an environmental justice inquiry in Baltimore, Maryland. *Ann. Assoc. Am. Geogr.* **99**, 767–787 (2009).
- Agyeman, J., Bullard, R. D. & Evans, B. *Just Sustainabilities: Development in an Unequal World* (MIT Press, 2003).
- Boone, C. G., Cadenasso, M. L., Grove, J. M., Schwarz, K. & Buckley, G. L. Landscape, vegetation characteristics, and group identity in an urban and suburban watershed: why the 60s matter. *Urban Ecosyst.* **13**, 255–271 (2010).
- Pellow, D. N. Toward a critical environmental justice studies: black lives matter as an environmental justice challenge. *Du Bois Rev. Soc. Sci. Res. Race*, <https://doi.org/10.1017/s1742058x1600014x> (2016).
- Roman, L. A. et al. Human and biophysical legacies shape contemporary urban forests: a literature synthesis. *Urban For. Urban Green* **31**, 157–168 (2018).
- Appel, I. & Nickerson, J. Pockets of poverty: the long-term effects of redlining. *Ssrn*, <https://doi.org/10.2139/ssrn.2852856> (2016).
- Hillier, A. E. Residential security maps and neighborhood appraisals: the home owners' loan corporation and the case of Philadelphia. *Soc. Sci. Hist.* **29**, 2 (2005).
- Rothstein, R. *The Color of Law: A Forgotten History of How Our Government Segregated America* (Liveright, 2017).
- Crossney, K. & Bartelt, D. The legacy of the home owners' loan corporation. *Hous. Policy Debate* **16**, 37–41 (2005).
- Jackson, K. T. Race ethnicity, and real estate appraisal: The Home Owners Loan Corporation and the Federal Housing Administration. *J. Urban Hist.* **6**, 419–452 (1980).
- Hillier, A. E. Redlining and the Home Owners' Loan Corporation. *J. Urban Hist.* <https://doi.org/10.1177/0096144203029004002> (2003).
- Woods, L. L. The federal home loan bank board, redlining, and the national proliferation of racial lending discrimination, 1921–1950. *J. Urban Hist.* <https://doi.org/10.1177/0096144211435126> (2012).
- Greer, J. The home owners' loan corporation and the development of the residential security maps. *J. Urban Hist.* <https://doi.org/10.1177/0096144212436724> (2013).
- Light, J. S. *The Nature of Cities: Ecological Visions and the American Urban Professions 1920–1960* (Johns Hopkins University Press, 2009).
- Lord, C. L. & Norquist, K. Cities as emergent systems: race as a rule in organized complexity. *Environ. Law* **40**, 2 (2010).
- US Commission on Civil Rights. *Report of the United States Commission of Civil Rights, 1959* (1959).
- Lipnitz, G. *The Possessive Investment in Whiteness: How White People Profit from Identity Politics* (Temple University Press, 2006).
- Gibson, C. & Jung, K. Historical census statistics on population totals by race, 1790 to 1990, and by hispanic origin, 1970 to 1990, for The United States, regions, divisions, and states. *Popul. Div.* **No. 76**, 1–116 (2005).
- Munnell, A. H., Tootell, G. M. B., Browne, L. E. & McEneaney, J. Mortgage lending in Boston: interpreting HMDA data. *Am. Econ. Rev.* <https://doi.org/10.2307/2118254> (1996).
- Pietila, A. *Not in My Neighborhood: How Bigotry Shaped A Great American City* (Ivan R. Dee, 2010).
- Hirsch, A. *Making the Second Ghetto: Race and Housing in Chicago 1940–1960* (University of Chicago Press, 1998).
- Satter, B. *Family Properties: How the Struggle over Race and Real Estate Transformed Chicago and Urban America* (Metropolitan Books, 2009).
- Woods, L. L. II "The inevitable products of racial segregation": multigenerational consequences of exclusionary housing policies on African Americans, 1910–1960. *Am. J. Econ. Sociol.* **77**, 967–1012 (2018).
- Kruse, K. & Sugrue, T. *The New Suburban History* (University of Chicago Press, 2006).
- Gee, G. C. A multilevel analysis of the relationship between institutional and individual racial discrimination and health status. *Am. J. Public Health* **92**, 615–623 (2002).
- Beyer, K. M. et al. New spatially continuous indices of redlining and racial bias in mortgage lending: links to survival after breast cancer diagnosis and implications for health disparities research. *Heal. Place* **40**, 34–44 (2016).
- Aaronson, D., Hartley, D. & Maxumder, B. *The Effects of the 1930s HOLC 'Redlining' Maps* (2017).
- Bullard, R. D. Solid waste sites and the black Houston community. *Sociol. Inq.* **53**, 273–288 (1983).
- Bullard, R. D. *The Quest for Environmental Justice: Human Rights and the Politics of Pollution* (Sierra Club Books, 2005).
- Bolin, B., Grineski, S. & Collins, T. The geography of despair: environmental racism the making of South Phoenix, Arizona, USA. *Human Ecol. Rev.* **12**, 156–168 (2005).
- Grineski, S., Bolin, B. & Boone, C. G. Criteria air pollution and marginalized populations: Environmental inequity in metropolitan phoenix, Arizona. *Soc. Sci. Q.* **88**, 535–554 (2007).
- Boone, C. G. in *Land of Sunshine* (eds Deverell, W. & Hise, G.) 167–178 (University of Pitts, 2005).
- Schlossberg, D. *Defining Environmental Justice: Theories, Movements, and Nature*. <https://doi.org/10.1093/acprof:oso/9780199286294.001.0001> (2007).

57. Boone, C. G. in *Sustainability for the 21st Century: Pathways, Programs, and Policies* (eds Hagen, B., Pijawka, D. & Hunt, K.) (2020).
58. Grove, J. M., Cadenasso, M. L., Pickett, S. T. A., Machlis, G. E. & Burch, W. R. J. *The Baltimore School of Urban Ecology: Space, Scale, and Time for the Study of Cities*. <https://doi.org/10.1017/CBO9781107415324.004> (2015).
59. Buckley, G. L. & Boone, C. G. To promote the material and moral welfare of the community: neighborhood Improvement Associations in Baltimore, Maryland, 1900–1945. In R. Rodger & G. Massard-Guilbaud (Eds), *Environmental and social justice in the city: Historical perspectives* 43–65 (White Horse Press, Cambridge, 2011).
60. Krimmel, J. *Persistence of Prejudice: Estimating the Long Term Effects of Redlining* (2017).
61. Bratman, G. N. et al. Nature and mental health: an ecosystem service perspective. *Sci. Adv.* **5**, eaax0903 (2019).
62. Hohl, B. B. C. et al. Creating safe and healthy neighborhoods with place-based violence interventions. *Health Aff.* **38**, 1687–1694 (2019).
63. Namin, S., Xu, W., Zhou, Y. & Beyer, K. The legacy of the Home Owners' Loan Corporation and the political ecology of urban trees and air pollution in the United States. *Soc. Sci. Med.* **246**, 112758 (2020).
64. Hoffman, J. S., Shandas, V. & Pendleton, N. The effects of historical housing policies on resident exposure to intra-urban heat: A study of 108 US urban areas. *Climate* **8**, 12 (2020).
65. Krieger, N. et al. Structural racism, historical redlining, and risk of preterm birth in New York City. *Am. J. Public Health* **110**, 2013–2017 (2020).
66. Hope, D. et al. Socioeconomics drive urban plant diversity. *Proc. Natl. Acad. Sci. USA* **100**, 8788–8792 (2003).
67. Hope, D. et al. Drivers of spatial variation in plant diversity across the Central Arizona-Phoenix ecosystem. *Soc. Nat. Resour.* **19**, 101–116 (2006).
68. Martin, C. A., Warren, P. S. & Kinzig, A. P. Neighborhood socioeconomic status is a useful predictor of perennial landscape vegetation in residential neighborhoods and embedded small parks of Phoenix, AZ. *Landsc. Urban Plan.* **69**, 355–368 (2004).
69. Leong, M., Dunn, R. R. & Trautwein, M. D. Biodiversity and socioeconomic in the city: a review of the luxury effect. *Biol. Lett.* **14**, 20180082 (2018).
70. Nier, C. L. Perpetuation of segregation: toward a new historical and legal interpretation of redlining under the Fair Housing Act. *John Marshall Law Sch. Law Rev* **32**, 617–665 (1999).
71. King, G., Keohane, R. O. & Verba, S. *Designing Social Inquiry*. <https://doi.org/10.1515/9781400821211> (2019).
72. Eisenman, T. S. et al. Urban trees, air quality, and asthma: an interdisciplinary review. *Landsc. Urban Plan.* **187**, 47–59 (2019).
73. Eisenman, T. S., Jariwala, S. P. & Lovasi, G. S. Urban trees and asthma: a call for epidemiological research. *Lancet Respir. Med.* **7**, e19–e20 (2019).
74. Rae, R. A., Simon, G. & Braden, J. Public reactions to new street tree planting. *Cities Environ.* **3**, 1–21 (2010).
75. Carmichael, C. E. & McDonough, M. H. The trouble with trees? Social and political dynamics of street tree-planting efforts in Detroit, Michigan, USA. *Urban For. Urban Green* **31**, 221–229 (2018).
76. Battaglia, M., Buckley, G. L. & Galvin, M. It's not easy going green: obstacles to tree-planting programs in East Baltimore. *Cities Environ.* **7**, 21 (2014).
77. Braverman, I. 'Everybody loves trees' Policing American cities through street trees. *Duke Environ. Law Policy Forum* **19**, 81–118 (2008).
78. Kirkpatrick, J. B., Davison, A. & Daniels, G. D. Resident attitudes towards trees influence the planting and removal of different types of trees in eastern Australian cities. *Landsc. Urban Plan.* **107**, 147–158 (2012).
79. Nelson, K. R. et al. in *American Panorama* (eds Nelson, R. K. & Ayers, E. L.) (2019).
80. Smith, M. L., Zhou, W., Cadenasso, M. L., Grove, J. M. & Band, L. E. Evaluation of the national land cover database for hydrologic applications in urban and suburban Baltimore, Maryland. *JAWRA J. Am. Water Resour. Assoc.* **46**, 429–442 (2010).
81. Grove, J. M., Locke, D. H., O'Neil-Dunne, J. P. M. & O'Neil-Dunne, J. P. M. An ecology of prestige in New York City: examining the relationships among population density, socio-economic status, group identity, and residential canopy cover. *Environ. Manage.* **54**, 402–419 (2014).
82. O'Neil-Dunne, J. P. M., MacFaden, S. W. & Royar, A. A versatile, production-oriented approach to high-resolution tree-canopy mapping in urban and suburban landscapes using GEOBIA and data fusion. *Remote Sens* **6**, 12837–12865 (2014).
83. MacFaden, S. W., O'Neil-Dunne, J. P. M., Royar, A. R., Lu, J. W. T. & Rundle, A. G. High-resolution tree canopy mapping for New York City using LIDAR and object-based image analysis. *J. Appl. Remote Sens.* **6**, 063567 (2012).
84. Taylor, P. et al. An object-based system for LiDAR data fusion and feature extraction. *Geocarto Int.* **28**, 1–16 (2012).
85. Core Team, R. R. *A Language and Environment or Statistical Computing* (2019).
86. Wickham, H. *tidyverse: Easily Install and Load the Tidyverse* (2017).
87. Pebesma, E. Simple features for R: standardized support for spatial vector data. *R J* **10**, 439–446 (2018).
88. Kassambara, A. *ggpubr: 'ggplot2' Based Publication Ready Plots* (2018).
89. Bates, D. M., Maechler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
90. Lüdtke, D. *sjPlot: Data Visualization for Statistics in Social Science*. <https://doi.org/10.5281/zenodo.1308157> (2018).
91. Lüdtke, D. *sjstats: Statistical Functions for Regression Models*. <https://doi.org/10.5281/zenodo.1284472> (2019).
92. Locke, D. H., Landry, S. M., Roy Chowdhury, R. & Chowdhury, R. R. What's scale got to do with it? Models for urban tree canopy. *J. Urban Ecol.* **2**, juw006 (2016).
93. O'Neil-Dunne, J. P. M. *A Report on the City of Baltimore's Existing and Possible Urban Tree Canopy* (2009).
94. Hox, J. J. *Applied Multilevel Analysis*. <https://doi.org/10.1017/cbo9780511610806> (1995).
95. Nakagawa, S. & Schielzeth, H. Coefficient of determination R^2 and intra-class correlation coefficient ICC from generalized linear mixed-effects models. *Ecol. Evol.* **14**, 20170213 (2017).
96. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach* (2nd ed). *Library of Congress Cataloging-in-Publication Data. Ecological Modelling* Vol. 172 (2002).
97. Nesbitt, L., Meitner, M. J., Girling, C., Sheppard, S. R. J. & Lu, Y. Who has access to urban vegetation? A spatial analysis of distributional green equity in 10 US cities. *Landsc. Urban Plan.* **181**, 51–79 (2019).
98. Locke, D. H. Residential housing segregation and urban tree canopy in 37 US cities; data in support of Locke et al. 2021 in npj Urban Sustainability ver 2. Environmental Data Initiative, <https://doi.org/10.6073/pasta/4ccbc7087959dc2a25063e589dee7718> (2020).

ACKNOWLEDGEMENTS

This work was supported by the National Socio-Environmental Synthesis Center (SESYNC) under funding received from the National Science Foundation DBI-1639145. This work was also supported by the Baltimore Ecosystem Study (BES) DEB-1637661 and DEB-1855277. Thanks to all data providers, the team in the Spatial Analysis Lab at the University of Vermont, and the Mapping Inequality project. Thanks to Mary Cadenasso and the Cadenasso Landscape and Urban Ecology lab at UC Davis for the Sacramento, CA data, and Monika Moskal and the Remote Sensing and Geospatial Analysis Laboratory at the University of Washington for the Seattle data, and NCDC Imaging and City and County of Denver GIS team. Thanks to Shawn Landry and the Water Institute at the University of South Florida for the Tampa, FL data and for providing thoughtful input on an earlier draft that improved the paper. Thank you Liz Wise, Brian Falasca, Michele Romolini, and LARIAC for access the LA Land cover data. Amanda Phillips de Lucas provided important insights about large-scale infrastructure projects and urban renewal. The findings and opinions reported here do not necessarily reflect those of the funders of this research.

AUTHOR CONTRIBUTIONS

D.H.L., B.H., and J.M.G. designed the research; D.H.L. and B.H. performed the research; D.H.L. and J.O.D. analyzed the data; D.H.L., B.H., J.M.G., S.T.A.P., L.A.O., C.F.A., and C.G.B. interpreted the data and findings; D.H.L., B.H., J.M.G., S.T.A.P., L.A.O., C.F.A., and C.G.B. revised and provided critically important content; and D.H.L., B.H., J.M.G., S.T.A.P., L.A.O., C.F.A., and C.G.B. wrote the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42949-021-00022-0>.

Correspondence and requests for materials should be addressed to D.H.L.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory

regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2021