Rethinking redlining: Environmental inequality within and between U.S. neighborhoods

January 27, 2026

Preliminary. Please do not cite or distribute.

Abstract

Environmental inequalities, such as unequal exposure to pollution and climate risks, persist across racial and socioeconomic groups in the United States. This paper re-examines the role of the Residential Security Maps created by the Home Owners' Loan Corporation (HOLC) in the 1930s, which graded neighborhoods according to perceived mortgage risk and have been widely linked to long-run racial segregation and environmental disadvantage. A common view holds that these maps not only reinforced residential segregation but also directly shaped the spatial distribution of environmental hazards, including air pollution, flood risk, and extreme heat. We evaluate this claim using a causal framework that combines machine-learning predictions of counterfactual HOLC grades in unmapped cities with a spatial difference-in-differences design. Our results confirm that the maps modestly increased racial sorting and segregation, consistent with prior work. However, we find no evidence that HOLC mapping independently affected the siting of environmental or climatic hazards. Differences in air pollution, flood risk, heat exposure, and mortality across historical grades are quantitatively similar in mapped and unmapped cities. These findings suggest that contemporary environmental inequalities primarily reflect residential sorting and discriminatory practices that operated broadly across U.S. cities, rather than an additional siting effect uniquely induced by the HOLC maps, which we do not detect.

Keywords: Air Pollution, Urban Economics, Economics of Minorities

JEL classification: Q53, R23, J15

1 Introduction

Environmental burdens—including pollution, climate risks, and associated health hazards—are unevenly distributed across geographic areas, racial groups, and income levels. A large body of work has shown that Black, Hispanic, and low-income populations in the United States are exposed to higher levels of fine particulate air pollution (Currie et al., 2023; Jbaily et al., 2022; Sager and Singer, 2025). Similarly, climate-related hazards disproportionately affect these communities, and climate change is likely to intensify such hazards over the coming decades. For example, urban heat islands have been found to affect non-white and lower-income populations disproportionately (Hsu et al., 2021; Benz and Burney, 2021), and flood risks are higher in historically marginalized neighborhoods (Wing et al., 2022). Yet, the origins of these correlations remain debated.

A frequent claim in policy and popular accounts is that the *Residential Security Maps* drawn in the 1930s by the Home Owners' Loan Corporation (HOLC) played a central role in shaping these disparities. As part of a federal mortgage initiative, HOLC produced color-coded maps for more than 200 U.S. cities to summarize the perceived financial risk of residential neighborhoods, assigning grades from A (lowest risk) to D (highest risk). These maps—later referred to as "redlining" maps—have been widely interpreted as having reduced credit access, reinforcing residential segregation and long-run neighborhood disadvantage. The policy has received substantial attention in the media (Plumer and Popovich, 2020; McCormick, 2022), among policymakers,[1] and in academic research. A growing body of studies outside economics documents strong correlations between historical HOLC risk grades and present-day environmental and health outcomes, while work exploiting local grade boundaries, such as Aaronson et al. (2021), has provided credible causal evidence of residential sorting. Much of the remaining literature, however, is descriptive and does not formalize an explicit counterfactual. In this paper, we contribute by proposing a complementary identification strategy that constructs a city-level counterfactual using unmapped cities around the HOLC population threshold, allowing us to reassess the causal effects of historical risk grades on contemporary environmental and health outcomes.

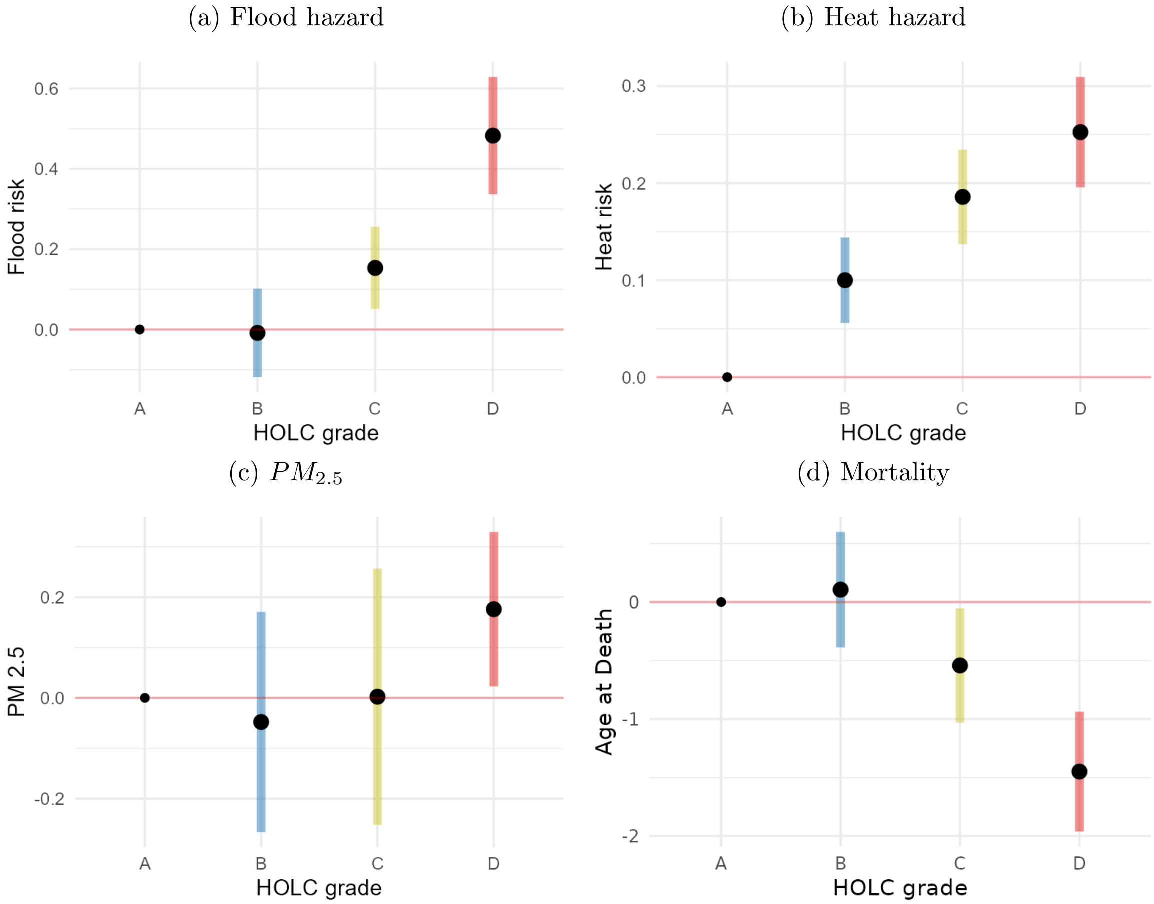
In this paper, we examine whether the documented correlation between historical neighborhood risk grades and contemporary environmental and health outcomes reflects a causal relationship. Figure 1 replicates the pattern that has been documented in several studies: the worse the neighborhood was graded in the 1930s, the higher the flood and heat hazard. Also, air pollution is higher in the worse graded neighborhoods. This pattern in environmental burden is also mirrored in mortality: the worse the historical risk grade, the lower the age at death.

This historical housing policy can affect environmental inequalites by sorting (households moving to areas with existing environmental hazards) or siting (the placement of environmental hazards in disadvantaged neighborhoods);[2] we investigate both of these channels. Our analysis

[1] For example, Senator Elizabeth Warren (D, MA) has proposed federal subsidies for first-time homebuyers in formerly redlined neighborhoods (see `https://elizabethwarren.com/plans/safe-affordable-housing`).

[2] Residential sorting leads to environmental disparities, as higher-income individuals are more willing to pay for amenities like clean air, resulting in wealthier communities with better environmental quality, while lower-income households tend to reside in more polluted areas. If income distributions differ across racial groups, this sorting mechanism can create a correlation between pollution exposure and race (Banzhaf and Walsh, 2008).

Figure (1) Environmental outcomes in 2020 across HOLC grades in treated cities



Notes: Standard errors are clustered at the city level. Bars reflect 95% confidence intervals.

has three core components. First, we use the 1930 full-count census to train a machine-learning model to predict which areas received different risk grades by the Home Owners' Loan Corporation (HOLC) and then use this model to predict the risk grade in areas of cities that were not mapped by HOLC to create a clean control group of comparable neighborhoods in untreated cities (Hynsjö and Perdoni, 2024). Second, we use a spatial difference-in-differences approach to show that the between-neighborhood differences across treated and untreated cities are balanced prior to the treatment. Third, we examine sorting and siting over the short and long run, combining census data since the 1940s with property-level climate hazard data, air pollution reanalysis data, high-resolution satellite imagery, and neighborhood-level mortality data to analyze disparities in pollution, climate, and health risks today.

Our results show that the *Residential Security Maps* modestly amplified racial sorting and segregation in treated cities. Consistent with earlier findings (Aaronson et al., 2021; Hynsjö and Perdoni, 2024), we document that the introduction of HOLC grades increased differential residential mobility across neighborhoods, primarily through higher out-migration of white households from lower-graded areas. These effects are economically meaningful but limited in magnitude, and they build on pre-existing differences across neighborhoods that were already present before the policy. Overall, the evidence suggests that the maps reinforced residential segregation patterns by coordinating housing market behavior rather than fundamentally reshaping urban demographics.

In contrast, we find little evidence that the policy had an independent effect on the siting of environmental and climatic hazards. While lower-graded neighborhoods are consistently exposed to higher levels of air pollution, flood risk, heat exposure, and worse mortality outcomes, these gradients are quantitatively similar in mapped and unmapped cities. Across a wide range of outcomes—including industrial activity, proximity to polluting highways, urban infrastructure, greenspace, pollution measures, climate hazards, and neighborhood-level mortality—we do not detect additional environmental burdens attributable to the HOLC maps themselves. Taken together, these findings indicate that the *Residential Security Maps* did not generate new patterns of environmental siting beyond those arising from more localized discriminatory practices and policies that operated in both treated and untreated cities.[3]

The paper is structured as follows: Section 2 provides background about the *Residential Security Maps* , Section 3 reviews the previous literature and describes the paper's contribution to the literature, Section 4 introduces the data and Section 5 describes our empirical strategy, and Section 6 discusses the results.

2 Historical background

In response to the mortgage crisis of the Great Depression, the Roosevelt Administration established the Home Owners' Loan Corporation (HOLC) in 1933 to refinance distressed mortgages on more favorable terms. By 1936, the agency had assisted over one million homeowners and held a substantial share of outstanding urban mortgages (Jackson, 1980). After completing its refinancing program, HOLC increasingly focused on assessing repayment risks and developing standardized appraisal techniques to stabilize property values and protect federal real estate investments (Woods, 2012; Michney, 2023). Policymakers and housing experts at the time emphasized that neighborhood conditions—rather than borrower characteristics alone—were central to the long-run performance of mortgage portfolios.

This shift in focus led to the launch of the *City Surveys* program in 1935. Facing time and resource constraints, HOLC confined the surveys to cities with populations above 40,000 and ultimately produced standardized *Residential Security Maps* and accompanying *Area Descriptions* for 239 cities. Although the program formally concluded in 1940, most maps were completed by mid-1938. Our empirical analysis exploits this population threshold and the resulting cross-city variation in map coverage as an exogenous historical shock.

To construct the *Residential Security Maps*, HOLC hired local real estate professionals and appraisal contractors to assess neighborhood-level lending risk (see Figure A1 for an example). Assessors grouped residential areas into one of four categories based on housing characteristics—such as age, quality, occupancy, and prices—as well as non-housing attributes, including race, ethnicity, and immigration status. Each neighborhood received a summary grade ranging from A (lowest perceived risk) to D (highest perceived risk), visualized using a color-coded scheme: green for A, blue for B, yellow for C, and red for D. This practice later became known as "redlining."

[3]Examples include zoning regulations (Shertzer et al., 2022), racial housing covenants (Sood and Ehrman-Solberg, 2023), and blockbusting.

Beyond summarizing housing conditions and demographics, the *Area Descriptions* frequently referenced environmental and land-use features believed to affect neighborhood stability. Contemporary appraisal manuals and professional guidelines treated proximity to parks, traffic corridors, and industrial activity as relevant determinants of residential desirability, often alongside explicitly racialized criteria. Green space was described as a protective buffer against commercial or industrial encroachment, while wide roads or parkways were viewed as barriers capable of containing the spread of "inharmonious uses." At the same time, industrial hazards and pollution were commonly discussed within the same evaluative framework as racial and ethnic "infiltration," reflecting prevailing ecological theories of neighborhood change.

These appraisal practices did not necessarily reflect an explicit federal intent to shape environmental outcomes. However, by codifying neighborhood desirability in ways that linked race, land use, and perceived environmental risk, the HOLC maps might have influenced how public and private actors understood appropriate locations for residential investment and, conversely, for industrial uses. As zoning decisions and siting choices responded to these classifications, low-graded neighborhoods may have become disproportionately exposed to environmental disamenities.

These links between appraisal, land use, and environmental conditions were not unique to HOLC. The Federal Housing Administration's *Underwriting Manual* explicitly classified the presence of racial and ethnic minorities alongside physical and environmental hazards when evaluating neighborhood risk. The manual described proximity to green space as a protective factor, noting that "a location close to a public park or area of similar nature is usually well protected from infiltration of business and lower social occupancy coming from that direction," while major roads or parkways were seen as barriers that could limit the expansion of "inharmonious uses." In contrast, neighborhoods lacking such buffers were viewed as more vulnerable to industrial encroachment. As a result, contemporaneous zoning and land-use regulations often operated to shield high-graded, predominantly white neighborhoods from polluting or environmentally hazardous activities, while implicitly channeling such uses toward lower-graded areas. California EPA (2021) documents how this regulatory configuration encouraged the relocation of polluting land uses into neighborhoods assigned unfavorable risk grades. Similarly, Rothstein (2017) argues that zoning laws "attempted to protect white neighborhoods from deterioration by ensuring that few industrial or environmentally unsafe businesses could locate in them," leaving polluting industries with limited options other than siting near African American communities. Taken together, these mechanisms suggest that HOLC's *Residential Security Maps*, even if not designed to govern environmental policy directly, could have contributed to the disproportionate concentration of industrial facilities and pollution in low-graded neighborhoods.

The broader economic impact of the *Residential Security Maps* depends on how widely they circulated beyond HOLC itself. Although the maps were formally classified as confidential, historical evidence suggests that their content and methods were far from obscure. Hundreds of consultants participated in their creation, thousands of local actors were contacted during the surveys, and HOLC officials reported sustained demand for access from practitioners familiar with the project (Michney, 2022). Selected maps were displayed at professional workshops, and the underlying appraisal techniques were promoted in trade journals (Winling and Michney, 2021). Rather than being fully secret or fully public, the maps functioned as an "open secret" within housing and real estate circles.

This interpretation is consistent with recent empirical evidence documenting persistent effects of red and yellow risk grades on racial segregation, economic opportunity, and intergenerational outcomes (Aaronson et al., 2021, 2023; Hynsjö and Perdoni, 2024). Yet, despite growing interest in the environmental dimensions of redlining, existing studies have not been able to credibly assess the causal impact of redlining on pollution exposure and environmental quality, largely due to the absence of appropriate control groups. Our empirical strategy is designed to address this limitation by leveraging quasi-experimental variation in HOLC map coverage to isolate the environmental consequences of this historically influential classification system.

3 Contributions to related literature

This analysis contributes to the literature in several ways. Previous studies on environmental inequality have focused on two broad roots for why marginalized communities bear a larger environmental burden.[4] It could be either due to household sorting or (dis)amenity siting. Our study provides a novel test of (dis)amenity siting but also contrasts this with the effects on residential sorting.

Numerous studies on household sorting demonstrate that environmental inequality can emerge as marginalized communities tend to move into more polluted areas. Or, conversely, that privileged communities move away from polluted areas. In a seminal study, Banzhaf and Walsh (2008) examine residential sorting due to the entry and exit of firms reporting to the Toxic Release Inventory (TRI)[5]. They find clear evidence that, subsequent to a firm entry, surrounding neighborhoods become poorer on average, leading the authors to conclude with the notion that "households vote with their feet". Depro et al. (2015) also find evidence for household sorting in Los Angeles due to differential cleanup efforts in face of a cancer risk. They point out that Hispanics have a lower marginal willingness to pay to avoid cancer risk. Bakkensen and Ma (2020) focus on residential sorting as a result of different flood risk in Florida. Based on house sales in a boundary discontinuity design, they find clear evidence of low-income and minority households sorting into higher flood risk zones. Christensen and Timmins (2022) examine how sorting takes place in practice. In a randomized experiment, testers with similar preferences and budget were assigned to an advertised listing by a real estate agent. Their findings reveal that minority testers were significantly more likely to be *steered* toward inferior neighborhoods that exhibit higher pollution levels among other characteristics. Heblich et al. (2021) examine the long-run impact of historical industry locations on neighborhood sorting in England. They find that during the Industrial Revolution in England, winds transported air pollution towards the East side of industrial cities, making these areas less attractive. This exogenous disamenity leads to residential sorting which increased economic deprivation in these affected areas. Beyond the high-income countries, Chen et al. (2022) examine the impact of changing air pollution in Chinese counties. They exploit 5-year average thermal inversions as IV that amplify air pollu-

[4]For reviews of research on environmental justice, including the mechanisms driving these disparities, see Banzhaf et al. (2019) and Cain et al. (2024).

[5]We also rely on this data source in the present study.

tion and find substantial migration outflows as a consequence. The effects are predominantly driven by younger, high-education individuals, which changes the local population composition. Anderson (2020) examines health outcomes along highways in Los Angeles. He finds that air pollution from traffic only occurs on the downwind side and finds that mortality is elevated by 3-5%. Crucially, he compares property prices across both sides of the highway and finds that they are balanced. This suggests that, in this specific case of pollution from LA highways, households are not aware of the role of the wind direction in diffusing air pollution. This finding reconciles with previous literature which shows that the new placement of salient urban disamenities, such as waste sites or factories, affects surrounding property prices (Farber, 1998; Schütt, 2021, for reviews). However, invisible hazards are often underestimated due to limited information (Barwick et al., 2024; Moulton et al., 2024). Hausman and Stolper (2021) shows that, even under uniformly limited information about pollution, large exposure differentials can arise through sorting due to the correlation of salient disamenities, such as smoke stacks, with hidden pollution. Overall, the literature has predominantly examined how the proportion of marginalized communities changes following the introduction of new pollution sources. In contrast, our research examines how siting decisions react in the aftermath of a sorting shock. We also examine whether the policy had repercussions for home affordability and home ownership that drive the initial sorting shock.

Studies on strategic siting of polluting factories find mixed evidence regarding the role of race and income in influencing siting decisions. Wolverton (2009) focuses on firms reporting to the TRI. She contrasts the local population characteristics long after the firm has been opened, common in the research field, with the characteristics at the time of the siting itself. The results show that long after the firm has opened, both race and income correlate with the firm location. In contrast, at the time of the siting, only income correlates with a firm location, while race is insignificant. This is consistent with no differential siting based on race, but, again, suggests that marginalized households sort into the hazardous neighborhoods. Silva et al. (2016) examines the local population composition at the time of entry and exit of US Texan firms reporting to the TRI. Controlling for factor prices and other variables influencing firm siting decisions, the authors find that polluting firms are more likely to locate into areas with higher share of nonwhites. Our study contributes to this line of research with a new test of disamenity siting decisions. Clay et al. (2025) examine fossil fuel siting and subsequent sorting in the US for the periodfind no evidence that fossil fuel plants were overly sited into African American neighborhoods.

We also contribute to the strand of research that has examined the role of historical redlining policies for today's environmental and health outcomes.[6] The recent digitization and open availability of HOLC maps has enabled a rapidly expanding body of quantitative research—primarily outside economics—documenting strong correlations between historical risk grades and contemporary environmental exposures and health outcomes, including air pollution, fossil fuel power plants, heat stress, tree canopy, biodiversity, and a range of morbidity and mortality measures (Lane et al., 2022; Locke et al., 2021; Nardone et al., 2020, 2021; Tessum et al., 2019; Namin et al., 2020; Hsu et al., 2021; Cushing et al., 2022). While this literature provides compelling

[6]Several studies investigate the legacy of redlining: Aaronson et al. (2021); Fishback et al. (2023); Hynsjö and Perdoni (2024). Miñano-Mañero (2024) studies how enhancing waterfronts and increasing tree canopy led to the upgrading of formerly redlined areas using geographic variation in tree plagues as an instrument.

descriptive evidence of persistent environmental inequality, it remains unclear whether these patterns admit a causal interpretation. Neighborhoods assigned different HOLC grades already differed systematically prior to the policy along dimensions related to housing quality, land use, and residential location, raising concerns that unaccounted-for pre-existing differences may bias estimates.

4 Data

HOLC maps. We incorporate HOLC grades using the digitized *Residential Security Maps* provided by the Mapping Inequality initiative, a project of the Digital Scholarship Lab at the University of Richmond (Nelson et al., 2023). The files contain maps for 216 cities across 39 states.[7] We convert the neighborhood polygons originally traced by HOLC into a regular grid of hexagonal cells, which constitute our primary spatial unit of observation. The use of a uniform grid facilitates comparison across space and simplifies the construction of counterfactual HOLC maps in control cities. Each hexagon has an area of 0.025 square kilometers (7.3 acres) and a side length of approximately 100 meters, a scale comparable to a typical city block in grid-plan cities such as New York City or Chicago.[8] A hexagon is assigned a HOLC grade if a single color occupies at least 75% of its surface.[9] This spatial transformation has a negligible impact on the overall distribution of grades (Hynsjö and Perdoni, 2024).

Historical census data. To study whether HOLC redlining constituted a shock to residential sorting and segregation, we assemble population data at the hexagon level from 1910 to 2010. For the period 1910–1940, we rely on geocoded full-count census records (Hynsjö and Perdoni, 2024). Heads of household are geocoded using street addresses from the proprietary census versions provided by *ancestry.com* and IPUMS (Ruggles et al., 2024). Addresses are cleaned following established practices in the spatial history literature (Logan and Zhang, 2018), and geographic coordinates are assigned using ESRI StreetMap Premium (2024), which combines parcel centroids and street locations. The resulting coordinates allow us to aggregate individual observations to our hexagon grid and construct neighborhood-level measures of demographic composition and housing characteristics. Consistent with prior work, HOLC grades reflect pre-existing socioeconomic patterns: racial composition, homeownership, property values, and rents vary systematically across grades (Hynsjö and Perdoni, 2024). After 1940, we rely on publicly available census data. We obtain tract-level measures for 1960–2010 from the National Historical Geographic Information System (NHGIS) (Manson et al., 2021). Census tracts are the smallest geographic units consistently identifiable over this period, but they are larger and more heterogeneous than our hexagons.[10] We convert tract-level data to the hexagon level by assigning each hexagon the characteristics of the census tract containing its centroid.

We exploit recently developed crosswalks linking individuals across census decades to study

[7] We use the version of the digitized maps released in December 2023. Relative to earlier releases used in the recent literature, this update includes twelve additional HOLC City Survey Maps (Nelson et al., 2023).

[8] The side length corresponds to roughly 328 feet.

[9] Our results are robust to modest variations in this threshold. Given the small size of each spatial unit, the majority of hexagons (80.9%) contain only one grade, while 7.6% do not meet the 75% criterion and are left ungraded.

[10] The median census tract covers an area equivalent to 34 hexagons.

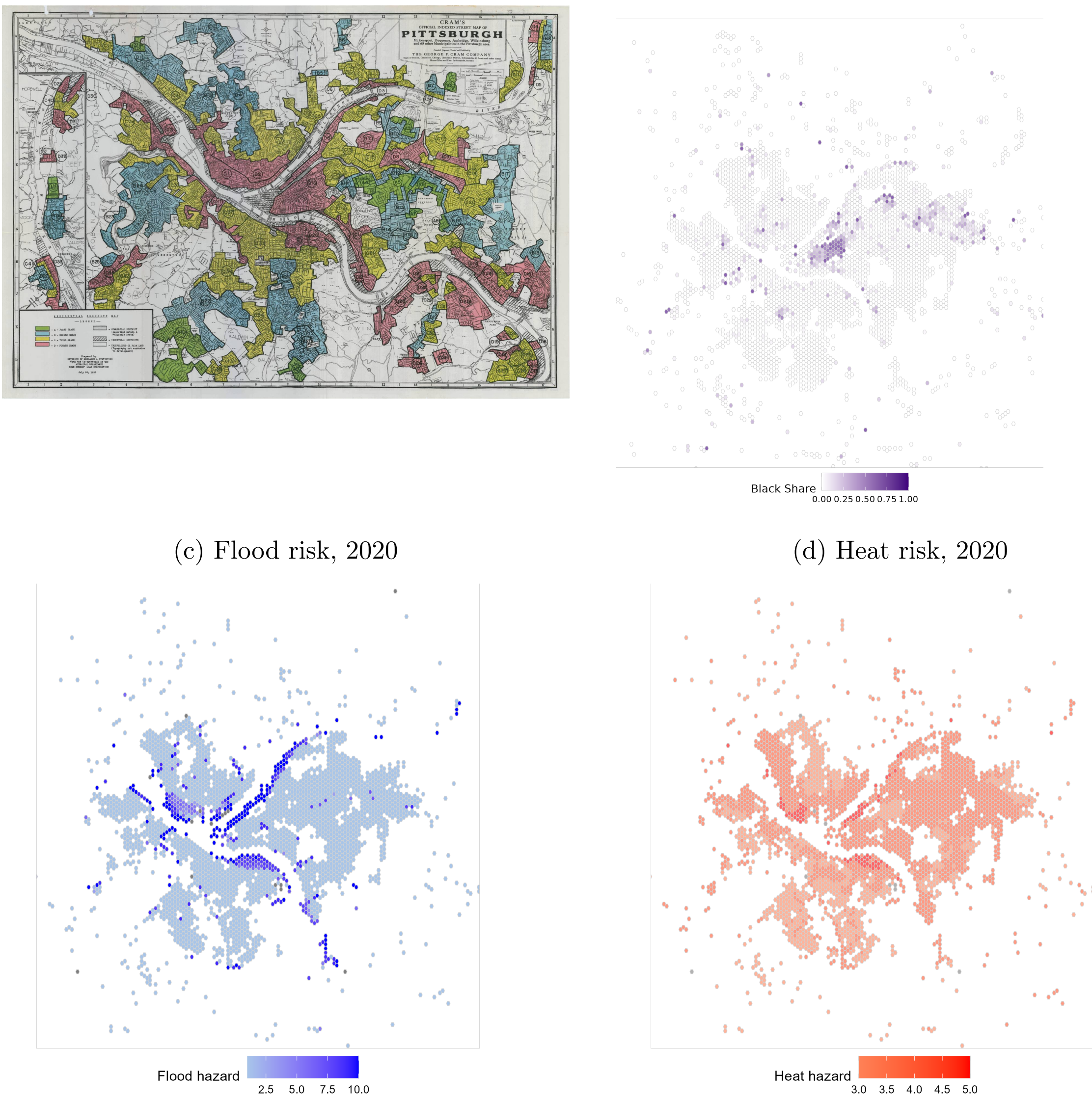
Figure (2) Example, Pittsburgh PA

(a) HOLC residential Security Map, 1937

(b) Black Share, 1930

(c) Flood risk, 2020

(d) Heat risk, 2020



Notes: Example of the data curation for Pittsburgh, PA: panel a) shows a digitized residential security map in 1937, panel b) shows the share of black population in the neighborhoods in 1930 according to the full-count census, panel c) and panel d) show the contemporaneous environmental and climatic risks of flood and heat in the neighborhoods in 2020.

residential mobility and longer-run individual outcomes. In particular, we use the IPUMS Multigenerational Longitudinal Panel (Ruggles et al., 2025) to compare individual residential locations between 1930 and 1950. The same linking initiative allows us to match census respondents to Social Security death records.[11] This information enables us to observe age and place of residence at death for individuals who lived in our cities of interest between 1930 and 1950, complementing our analysis of residential sorting with evidence on subsequent mobility and persistence.

Environmental and climatic hazards. Our preferred measure of environmental and climatic hazards—flood, heat, and fire risk—comes from a proprietary dataset by First Street Foundation (2022). This dataset provides property-level risk assessments on a scale from 1

[11]The death records are public and disseminated by the National Archives and Records Administration. A researcher-friendly version is provided by Hollingsworth et al. (2024).

(low risk) to 10 (high risk) for 142 million properties across the US The risks are climate-adjusted, reflecting 30-year projections based on the IPCC CMIP5 RCP4.5 model. The data integrates publicly available inputs at 30m resolution (e.g., satellite imagery, temperature data) with proprietary inputs to create address-level hazard estimates. For example, Figure 2, panel c) and d) illustrate flood risk and heat risk aggregated from the property-level to our hexagon grid, which is our unit of analysis. Light blue (red) shading indicates low risk and dark blue (red) shading indicating high risk. The measure combines physical risk with information on local mitigation and adaptation efforts, including "gray" infrastructure like levees and pumps, and "green" solutions like wetland restoration.

Air Pollution. As two other dimensions of environmental hazard, we leverage gridded air pollution data that are typically emitted from combustion processes like industrial production, traffic or energy generation. To examine a widely used measure of air pollution, particulate matter $PM_{2.5}$ concentration, we link average grid values from Shen et al. (2024) to our hexagon grid. The $PM_{2.5}$ values come from a reanalysis project that combines ground-station and satellite measurements and a chemical transport model. It encompasses the whole US at 0.01° x 0.01° (ca. 1km x 1km) resolution between 1998-2022. For nitrogen dioxide NO_2 concentrations, we rely on the satellite product Copernicus Sentinel-5 Precursor TROPOspheric Monitoring Instrument (TROPOMI).[12] Is is delivered at resolution of 0.01° x 0.01° between 2019-2022. This data is reprocessed by the remote sensing engineers to ensure data accuracy. NO_2 concentrations are expressed in units of $Pmolec/cm2$ which we scale by $1e6$ for easier readability.

Siting of firms and urban disamenities. We unpack these siting results further by using data from the EPA's *Toxic Release Inventory* (TRI), which includes geocoordinates and annual emission details from 1987 to 2023. In 2010, 21,810 facilities reported toxic pollutant emissions. We define exposure as areas within a 300m radius of an industrial site, acknowledging that pollutant exposure decreases with distance but can extend up to 1,600m (Currie et al., 2015).

In addition, we use the *Global Human Settlement Layers* (GHSL), which classify human land use at a 10m resolution into 15 categories, including vegetation, roads, residential, and commercial/industrial areas. From these data, we compute the share of each land use type within a neighborhood.

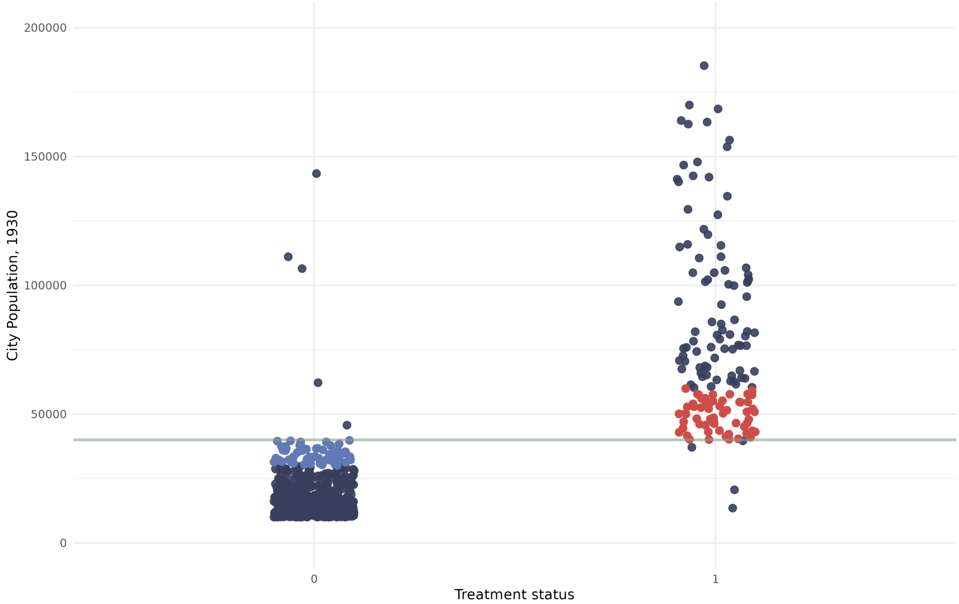
5 Empirical strategy

Our identification strategy builds on the design proposed by Hynsjö and Perdoni (2024), who exploit the population threshold determining whether a city was surveyed by the HOLC. While their approach relies on a temporal difference-in-differences framework, we adopt a spatial difference-in-differences design that replaces the time dimension with cross-sectional grade variation. This choice is dictated by data availability, as measures of pollution and climate hazards are not observed prior to the introduction of the HOLC maps. As in Hynsjö and Perdoni (2024), identification hinges on the fact that only cities with populations above 40,000 residents in 1930 were mapped. We therefore compare neighborhoods in mapped cities to observationally similar neighborhoods in unmapped cities just below the threshold. To ensure comparability around

[12] We use *Offline tropospheric NO_2 column number density* as common in the literature.

the cutoff, we restrict attention to cities with populations between 30,000 and 40,000 residents (control group) and between 40,000 and 60,000 residents (treated group) in 1930, resulting in a near-symmetric sample of 48 control and 51 treated cities, as shown in Figure 3.[13]

Figure (3) Population Threshold for HOLC City Surveys



Notes: The graph shows the treatment status of cities according to their 1930 population. The horizontal green line highlights the 40,000 people threshold. Blue points identify cities in the control group, while red ones highlight treated cities.

5.1 Constructing Counterfactual HOLC Maps

To construct a credible counterfactual for unmapped cities, we use a machine-learning classification approach that replicates the HOLC grading process using pre-treatment data only. Specifically, we train a random forest model on neighborhoods located in mapped cities, using information from the 1930 full-count census (Ruggles et al., 2024), and then apply the trained model to predict counterfactual HOLC grades in unmapped cities. The algorithm is thus used to replicate HOLC appraisal decisions rather than to forecast post-treatment outcomes. By construction, the training procedure excludes all post-treatment information, and control-city neighborhoods are never used in model estimation.

Our approach can be interpreted as a matching strategy that provides an alternative to synthetic control methods. Rather than constructing a weighted average of control units to mirror treated units before the intervention, we classify neighborhoods according to predicted HOLC grades. For each observed grade in treated cities, the corresponding control group consists of neighborhoods in unmapped cities that the model predicts would have received the same grade had they been surveyed. This procedure yields transparent and easily interpretable donor pools for each grade, which can be visualized spatially, and shifts the burden of validity to out-of-sample checks—such as parallel trends—that are not explicitly targeted during model training.

[13]To limit potential spillovers, we further restrict the control group to cities located at least 30 kilometers from the nearest mapped city.

The unit of analysis is a regular grid of hexagons, each with a side length of 100 meters and an area of 0.025 km^2, approximating the size of a typical city block. HOLC grades are taken from the digitized *Residential Security Maps* provided by the Mapping Inequality initiative (Nelson et al., 2023). A hexagon is assigned a grade if a single color covers at least 75% of its surface. The classification model is trained on 47 socioeconomic and housing variables from the 1930 census, measured at multiple geographic scales—hexagon, surrounding buffers, city, and county—yielding a total of 158 predictors. The construction of these measures relies on the geocoded census data described in Hynsjö and Perdoni (2024).

We implement a random forest classifier (Breiman, 2001), which is well suited to this setting given the nonlinearities, interactions, and class imbalance inherent in the HOLC grading problem. Model performance is assessed using out-of-sample validation on a held-out test sample. The confusion matrix reported in Appendix Table **??** shows an overall accuracy exceeding 90%, with class-specific sensitivities above 90% for C and D grades. Importantly, the predicted grade distribution closely mirrors the observed distribution, indicating that the algorithm does not mechanically distort the prevalence of risk categories. Performance remains similarly high when restricting the test sample to smaller cities, which are most relevant for our identification strategy (Hynsjö and Perdoni, 2024).

Beyond predictive accuracy, the spatial structure of the predicted maps closely resembles that of the original HOLC maps. Figure 4 compares observed and predicted grades for Pittsburgh, PA, illustrating that the model reproduces compact and contiguous neighborhood boundaries rather than fragmented or noisy patterns.[14] Finally, we assess the drivers of the classification using variable importance measures in Appendix Figure A3. Socioeconomic status and housing market characteristics contribute most to predictive power overall, followed by racial composition. When focusing on the lowest grade, racial composition becomes relatively more important, consistent with historical accounts of HOLC appraisal practices. At the same time, no single variable category dominates the model, underscoring that the algorithm captures the multidimensional nature of the HOLC grading process rather than relying on any single input.

5.2 Differences-in-differences

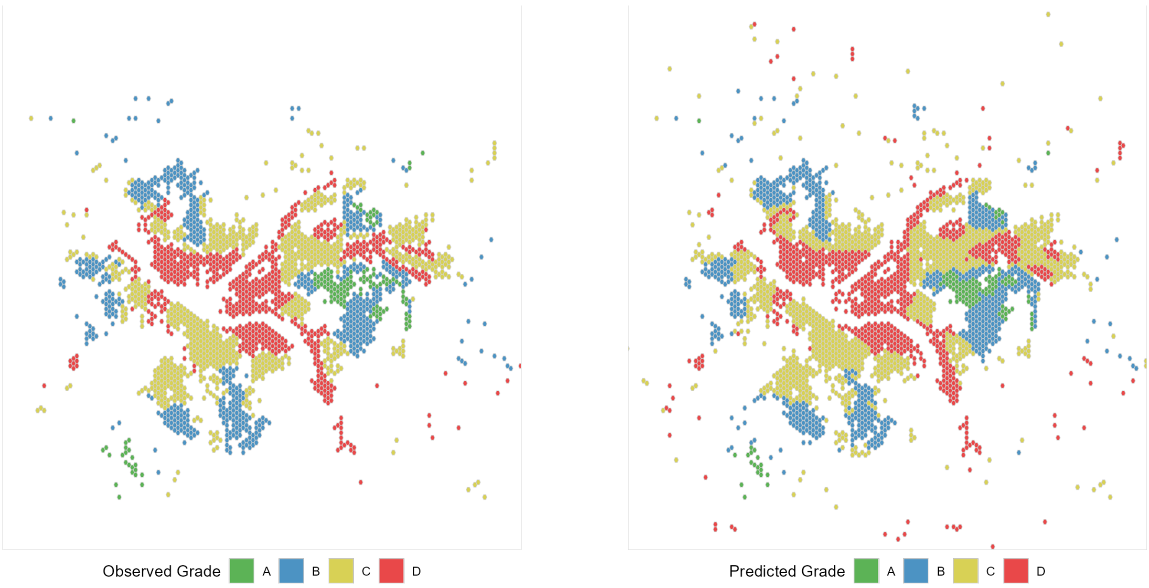
Our estimation compares outcomes between neighborhoods with different grades within treated cities to differences between neighborhoods that would have received the same grade within counterfactual (unmapped) cities. In our baseline specifications, we consistently use predicted grades to categorize hexagons both in the control and in the treatment group, in order to mitigate any systematic bias due to remaining idiosyncrasies that could stem from the grades assigned by the random forest in comparison cities in contrast to the true grades in treated cities.[15] Thus, our spatial difference-in-differences approach is based on pair-wise comparisons between predicted grades (D vs. C, C vs. B, or D vs. B) and, in effect, utilizing this within-city grade

[14] Predicted maps for additional treated and control cities display similar coherence, supporting the use of the classification algorithm as a tool for reconstructing plausible HOLC-style maps in cities that were never surveyed. A complete collection of predicted maps can be accessed at this link: *Predicted Maps*

[15] To examine the sensitivity of our results to this choice, we also perform the analysis with observed grades in treated cities *versus* predicted grades in comparison cities, with unchanges results.

Figure (4) Comparing observed vs. predicted grade in Pittsburgh, PA

(a) Observed grades (b) Predicted grades



Notes: The figure compares the original digitized HOLC map for Pittsburgh, PA (Nelson et al., 2023) with the corresponding neighborhood classifications generated by our random forest model at the hexagon level. Colors indicate HOLC grades, with green denoting A, blue B, yellow C, and red D. Predicted maps for all cities in the sample are available at *Predicted Maps*.

difference instead of a time dimension. We run the following regression equation:

$$y_{ic} = \alpha LowGrade_i + \gamma Treated_c + \beta\,(LowGrade_i \times Treated_c) + \delta_c + \mathbf{z}_i'\boldsymbol{\phi} + e_{ic}$$

where y_{ic} is an outcome, such as African American share or environmental hazard, in a neighborhood hexagon i, in city c, $LowGrade_i$ indicates whether a hexagon is the lower (predicted) grade among HOLC grades —i.e., D among D-C comparisons or D-B comparisons, C among C-B comparisons—, $Treated_c = \mathbb{I}(pop > 40k)$ indicates whether city c was treated, δ_c are fixed effects for cities, $\mathbf{z}_i$ are pre-determined neighborhood-level controls which may be optionally included. Our coefficient of interest is the interaction of the two indicators $LowGrade_i \times Treated_c$, which reflects the additional effect of being in a lower-graded neighborhood in the treated cities relative to the lower-graded neighborhood in the control cities. The units are weighted by 1930 neighborhood population, and standard errors are clustered at the city level.

6 Results

6.1 Balancing and validation

We first validate whether neighborhood disparities were balanced before the policy in 1930 (after standardizing all variables). Table 1, panel (a), compares the neighborhood inequalities for mapped cities, relative to comparison cities. The findings show that in comparison cities, D-graded neighborhoods were more precarious and more segregated than C-graded neighborhoods, with lower home ownership, lower house values, lower rent prices and a higher Black share. The coefficient on $LowGrade_i \times Treated_c$, however, indicates no statistically significant additional differences between D- and C-neighborhoods in treated cities, suggesting disparities were similar across treated and untreated cities before the policy.[16]

Table (1) Balance of neighborhood characteristics on D-C

Sample:	1930			
	Black share (1)	Ownership rate (2)	House value (3)	Rent price (4)
LowGrade × Treated	0.01	0.00	0.06	0.15
	(0.21)	(0.10)	(0.10)	(0.14)
LowGrade	0.62***	-0.38***	-0.49***	-0.27*
	(0.14)	(0.06)	(0.07)	(0.14)
Dependent variable mean	0.10	-0.13	-0.27	-0.06
Observations	18,665	18,657	18,615	18,615
City fixed effects	✓	✓	✓	✓

Census Year:	1920		1910	
	Black share (1)	Ownership rate (2)	Black share (3)	Ownership rate (4)
LowGrade × Treated	-0.08	0.04	-0.09	-0.02
	(0.20)	(0.09)	(0.17)	(0.07)
LowGrade	0.51***	-0.42***	0.42***	-0.36***
	(0.13)	(0.06)	(0.11)	(0.05)
Dependent variable mean	0.09	-0.09	0.08	-0.06
Observations	15,582	15,577	13,428	13,421
City f.e.	✓	✓	✓	✓

Notes: Regression of neighborhood characteristics in 1930, 1920, 1910, prior to redlining. The observations are weighted by 1930 neighborhood population. Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01

Panel (b) of Table 1 extend this balance test to the census years 1910 and 1920 for those outcomes that are consistently obtained. Again, we find substantial differences between neighborhoods: higher prevalence of African American communities and substantially lower home ownership rate. And, again, any additional difference in treated cities are absent. In sum, neighborhoods which would later be labeled low-grade, already had worse conditions in 1930, 1920 and 1910, but disparities were comparable between treated and untreated cities. This could be interpreted as the result of other local forms of discrimination, happening before our treatment, *both* in treated cities and untreated cities. These balance tests highlight a significant challenge in the literature: Previous studies document consistently large imbalances on different sides of a geographical grade boundary within treated cities (Aaronson et al., 2021; Fishback et al., 2023). Addressing pre-existing differences (and trends in these differences) is difficult without a counterfactual from unmapped cities: Aaronson et al. (2021), resort to focusing on

[16]Due to a still pending data request at the US Census for the 1930 census data, this balance test includes only cities within a symmetric bandwidth around the population cutoff, that is cities with population of 30-50 thousand. This results in 70 cities. For the remainder, we are able to examine all 99 cities between 30-60 thousand, where 51 are treated and 48 serve as comparison.

the fraction of boundaries that satisfy balance using propensity score methods. However, this might forgo a wide share of variation. By contrast, our approach allows to compare entire neighborhoods, while acknowledging the pre-treatment difference between grades, using unmapped cities as counterfactual.

To corroborate the comparability of neighborhoods, we, next, perform the balance test for C vs. B neighborhoods, finding similar patterns. Table A2 shows that lower-graded neighborhoods had lower home ownership, lower rents and house values and were more segregated across both treated and comparison cities, with no significant differences for treated cities at the 95% level.

6.2 Sorting and housing prices

We now turn to the effects of the *Residential Security Maps* on individual residential mobility. Table 2 presents results whether individuals moved out of their 1930 residence, proxied by their zip-code. Columns 1 to 3 focus on short-term mobility between 1930 to 1950, while column 4 to 6 focus on mobility from 1930 until death: Overall, living in a low graded neighborhood, D vs. C, is associated with lower mobility. Individuals are less likely to move by 5 percentage points. This can be interpreted that marginalized individuals are "locked in" to their neighborhoods. However, this differs for individuals who live in formerly redlined cities. The coefficient on *LowGrade x Treated* is positive and significant in column 1 and 4. Living in a redlined city leads to 2 percentage point higher mobility. Distinguishing between whites and blacks, the entire effect seems to be driven by white inhabitant moving out of low graded neighborhoods. By contrary, black inhabitants stay in the impoverished locations, which increases racial segregation.

Table (2) Sorting

Sample:	Short-term sorting (1950)			Lifetime sorting		
	All (1)	White (2)	Black (3)	All (4)	White (5)	Black (6)
LowGrade × Treated	0.022^{*}	0.028^{**}	-0.069	0.023^{**}	0.027^{***}	-0.041
	(0.013)	(0.012)	(0.044)	(0.009)	(0.010)	(0.022)
LowGrade	-0.050^{***}	-0.042^{***}	-0.143^{***}	-0.051^{***}	-0.046^{***}	-0.118^{***}
	(0.008)	(0.008)	(0.034)	(0.007)	(0.008)	(0.030)
Observations	691,988	664,429	27,559	204,880	191,643	13,237
City F.E.	✓	✓	✓	✓	✓	✓
Birth Year F.E.				✓	✓	✓

Notes: The table reports estimates of residential sorting across predicted HOLC grades in mapped (treated) and unmapped (control) cities. The unit of observation is an individual. *LowGrade* is an indicator for residence in a lower predicted HOLC grade within the relevant comparison (grade D versus C). *Treated* equals one for cities surveyed by the HOLC and zero for comparison cities below the population cutoff. Columns (1)–(3) measure short-term sorting using residential moves between 1930 and 1950. Columns (4)–(6) measure lifetime sorting by comparing residential location in 1930 to location at death, using Social Security death records. All specifications include city fixed effects; columns (4)–(6) additionally include birth-year fixed effects. Standard errors are clustered at the city level. *, **, and *** denote statistical significance at the 10, 5, and 1 percent levels.

The analysis on residential sorting thus far has focused on African American segregation. While African American households have lower income on average, it is unclear whether the residential sorting also affects low-income households irrespective of their ethnic group status. In the absence of households' income records, we resort to census information on the prevalence of home ownership reflective of household wealth, and the rent prices, reflective of affordability and quality of housing. The corresponding regressions are shown in Table A5 for the share of home ownership and in Table A4 for rent prices, each for an early period suggestive of short-term effects (1940) and for the latest available period suggestive of a long-term effect. Home ownership is consistently lower in lower graded neighborhoods. For instance, it differs by -6.7% between D- relative to C-graded neighborhoods. Treated cities are not notably different, as the coefficients on LowGrade x Treated indicate: All coefficients are statistically insignificant. Regarding rent a similar picture emerges: Rents are lower in worse graded neighborhoods, although this is statistically significant in 1990, but not in 1940. Treated cities do not differ systematically. Overall, this suggests: First, the wealth and affordability does not deteriorate noticeably due to redlining. But, second, the absence of statistically significant effects for income-related neighborhood characteristics suggests that the residential sorting is less driven by an income-gradient of households. Rather, the sorting decision of households is predominantly driven by their ethnic group status. In other words, households do not just move into or out of formerly redlined neighborhoods because of the mere housing conditions, which we find are hardly affected.

6.3 Siting and environmental exposures

Having established that the practice of redlining was a shock to a neighborhood's ethnic composition, we now focus on the siting of (dis-)amenities. Table 3 reports the findings within our estimation framework, which uses spatial differences within cities and compares these across cities near the population cutoff that differ in their treatment status. The findings show little evidence that disamenities are disproportionately sited in treated cities.

Panel a) tests for disproportionate siting of industry (column 1 to 3) and highways (column 4 to 6). If the *Residential Security Maps* facilitated firms to place their factories strategically close to marginalized communities, we would observe elevated presence of toxic plants in lower graded neighborhoods of treated cities. This is not supported by our findings. The estimates shows that, overall, lower-graded neighborhoods are more likely to have industrial areas (measured by high-resolution Land Use and Land Classification maps) or industrial toxic plants. However, this difference for mapped cities, reflected by the coefficient on *LowGrade x Treated* is never statistically significant at the 95% level and often trends in the opposite direction. This finding is consistent whether for different time periods from early industry measures in 1970s (general differences +0.04 more likely to have industry vs. additional difference in redlined cities: -0.02), as well as later industry measures up to 2010.

Another major environmental disamenity are highways. In Table 3, columns 4 to 6, the results exhibit more highway presence in lower graded neighborhoods. The redlined D neighborhoods are slightly more likely to have highway presence, as compared to yellow-lined C neighborhoods. However, when comparing this first difference to the additional difference in treated cities, the coefficient is statistically insignificant. This suggests that treated cities did not have additional highway construction through the lower graded neighborhoods. As a result, we cannot conclude that the *Residential Security Maps* amplified hazards through highway construction.

In panel b) of Table 3, we examine the siting of further urban disamenities. The findings point toward the same conclusion. In the 1970s, low-graded neighborhoods had higher commercial building shares and a lower residential building share (column 1 and 2). The earliest measures of greenness in 1985 point toward the fact that low graded neighborhoods had less vegetation (column 3) which can absorb excessive rainfall, and diffuse heat and air pollution, sometimes through fresh air corridors. Neither of these neighborhood disamenities is elevated in treated cities, in contrast to the baseline comparison cities. These findings hold even when turning to later outcomes measured using satellites in 2010 such as high-intensity urban areas, urban built-up and greenness (columns 4 to 6).

Next, in panel c), we turn to environmental and climatic hazards. When examining air pollution across neighborhoods in the two sets of cities, we find very similar evidence. Table 3 shows the double-difference estimates for particulate matter $PM_{2.5}$ (column 1 to 3) and nitrogen dioxide NO_2 (column 4).[17] The estimates for the grade differences within both treated and untreated cities are strong and statistically significant, suggesting that living in a low-grade neighborhood comes with substantially higher air pollution. This difference, however, is similar in treated cities: The double-difference estimate is again very close to zero. These findings suggests that air pollution hazards are concentrated in more precarious and more segregated neighborhoods. But, evenly so, both in treated cities as in untreated comparison cities.

Columns 5 and 6 extend to high-resolution address-level climatic hazards, such as flooding and extreme heat.[18] Again, low-graded neighborhoods bear a higher flood risk (+0.06) and a higher heat risk (+0.27), both of which are statistically significant at the 5- or 10-percent level. Mapped cities do not differ systematically with small coefficients in the case of flood, and wide confidence intervals in the case of heat.

The findings up to here suggest that while the policy has amplified residential sorting, the environmental hazards affect precarious households similarly no matter if a city was mapped by the HOLC or remained unmapped. No matter which proxy for disproportionate pollutant siting we use, we find evidence of the same pattern. The lower-graded side in each pair consistently exhibits more environmental disamenities: more toxic plants, more highway, more commercial areas, fewer residential areas and less vegetation cover. However, we rarely observe significant differences between treated cities and the counterfactual differences in untreated cities.

[17]Both pollutants are the result from combustion processes such as industry, traffic, electricity generation and vast evidence that shows a clear link to health outcomes. As described in **??**, the air pollution data are from a very recent reanalysis effort with both wide coverage and an unprecedented resolution, but the resolution is still coarse as compared to our neighborhood hexagons which leads to a high R^2 in the estimation.

[18]These hazards capture not only the geophysical core risk, but also risks as a result of human decision-making such as where to build protective infrastructure like levees, pumps or greenspace to absorb heat and excess rainfall

Table (3) Siting

Panel a) Firms and highways

Outcome:	Industry			Highways		
	1970s	1990	2010	1970	1990	2005
	(1)	(2)	(3)	(4)	(5)	(6)
LowGrade × Treated	-0.02	-0.03	0.01	0.00	0.00	0.01
	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.04)
LowGrade	0.04***	0.08***	0.03**	0.01*	0.01**	0.05***
	(0.01)	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)
Data	LULC	TRI	TRI	Highways	Highways	(Inter-)State
Dependent variable mean	0.04	0.08	0.03	0.01	0.02	0.25
Observations	30,850	30,845	30,845	30,850	30,845	30,845
City fixed effects	✓	✓	✓	✓	✓	✓

Panel b) Land use and greenness

Period:	Early			Later		
Year:	1970		1985	2010		
	Commercial	Residential	Greenness	High-intensity	Built-up	Greeness
	(1)	(2)	(3)	(4)	(5)	(6)
LowGrade × Treated	0.03	-0.01	0.00	0.01	0.01	-0.01
	(0.04)	(0.04)	(0.01)	(0.03)	(0.02)	(0.01)
LowGrade	0.05*	-0.11***	-0.03***	0.06**	0.04***	-0.02**
	(0.02)	(0.03)	(0.01)	(0.03)	(0.01)	(0.01)
Data	LULC	LULC	Satellite	LULC	Satellite	Satellite
Dependent variable mean	0.22	0.63	0.32	0.16	-0.46	0.34
Observations	30,850	30,850	30,845	30,845	30,845	30,845
City fixed effects	✓	✓	✓	✓	✓	✓

Panel c) Environmental and climatic hazards

Outcome:	Pollution				Climatic	
Hazard:	$PM_{2.5}$			NO_2	Flood	Heat
Year:	2000	2010	2020			
	(1)	(2)	(3)	(4)	(5)	(6)
LowGrade × Treated	0.09	-0.01	-0.01	0.00	0.00	-0.15
	(0.17)	(0.03)	(0.03)	(0.03)	(0.04)	(0.18)
LowGrade	0.33***	0.05*	0.06**	0.08***	0.06**	0.27*
	(0.11)	(0.03)	(0.03)	(0.03)	(0.03)	(0.15)
Dependent variable mean	1.68	4.58	13.15	9.81	7.78	38.89
Observations	30,268	30,268	30,845	30,845	30,845	30,845
City fixed effects	✓	✓	✓	✓	✓	✓

Notes: Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01

6.4 Mortality

Our analysis thus far has focused on specific environmental variables, such as proximity to toxic plants, vegetation cover, and land use patterns, to assess the siting of pollutants. However, these variables capture only a subset of potential environmental hazards, leaving open the possibility that other unobserved pollutants or cumulative exposures contribute to long-term disparities. Given the well-documented link between pollution and adverse health outcomes, including higher rates of respiratory and cardiovascular diseases, our final measure considers mortality. To this end, we link individuals residing in redlined areas during the 1930s to the Social Security death records. This data enables us to examine whether individuals from historically redlined neighborhoods exhibit higher mortality rates, capturing the cumulative health effects of exposure to disadvantaged living conditions over time.

Table (4) Mortality

Sample:	Age-at-death		
	(1)	(2)	(3)
LowGrade × Treated	-0.022	-0.052	-0.020
	(0.068)	(0.147)	(0.066)
LowGrade	-0.166***	-0.219**	-0.115**
	(0.054)	(0.097)	(0.051)
LowGrade × Treated × Moved		0.009	
		(0.150)	
LowGrade × Moved		0.058	
		(0.093)	
LowGrade × Treated × Black			0.104
			(0.109)
LowGrade × Black			0.034
			(0.131)
Observations	239,688	204,108	239,679
City F.E.	✓	✓	✓
Birth Year F.E.	✓	✓	✓

Notes: The table reports estimates for age at death (in years) using linked Social Security death records. The unit of observation is an individual. The sample includes individuals residing in neighborhoods (hexagons) with predicted HOLC grades and treatment status defined at the city level. *Treated* equals one for cities surveyed by the HOLC and zero for comparison cities below the population cutoff. *LowGrade* is an indicator for the lower grade within the relevant comparison; in the specifications shown it equals one for grade D and zero for grade C (predicted HOLC grades). All specifications are estimated by OLS with city fixed effects and birth-year fixed effects, and standard errors are clustered at the city level. Column (1) reports baseline estimates for the full sample. Column (2) allows the effect to vary by residential mobility, where *Moved* is an indicator equal to one if the individual's ZIP code at death differs from the 1940 ZIP code. Column (3) allows the effect to vary by race, where *Black* is an indicator for Black individuals. *, **, and *** denote significance at the 10, 5, and 1 percent levels.

Table 4 reports results for age-at-death where we residualize the data using birth year fixed effects. Our main results are also confirmed by the mortality data. People who resided on the D-graded side (or counterfactual D-graded side) died about 0.16 years (≈ 2 months) earlier on average. Yet, the interaction effect is not statistically significant. Column 2 includes further interactions with moved status to test whether individuals who were born in low-graded and redlined neighborhoods were are to compensate for this disadvantage by migrating out. Yet, the coefficient on the triple interaction implies that even those individuals who moved die at a similar younger age. Lastly, column 3 tests whether Black individuals had a differential impact. But, also here, the interaction on age at death is not statistically significant.

7 Conclusion

In the second half of the 1930s, the Home Owners' Loan Corporation undertook an unprecedented effort to survey and classify residential neighborhoods across more than 200 U.S. cities. Conceived as a data-driven initiative to standardize real estate appraisal and stabilize housing markets during the recovery from the Great Depression, the resulting *Residential Security Maps* quickly became an influential information tool for housing professionals. In the decades since, these maps have also come to symbolize institutionalized discrimination in public debate and policy discussions. This historical reassessment is grounded in both archival evidence and post–civil rights norms of equality, and it remains valid independently of the magnitude of any estimated causal effects. At the same time, careful quantitative analysis can still play a useful role by clarifying how, and through which channels, this widely debated federal intervention shaped patterns of urban inequality.

A central empirical challenge in this context is that HOLC grades were deliberately assigned along neighborhood boundaries that closely tracked pre-existing socioeconomic conditions. To address this challenge, we depart from spatial regression discontinuity approaches and instead exploit an exogenous population threshold that determined which cities were surveyed. By combining this design with a machine-learning classification model that reconstructs counterfactual HOLC maps in comparable, unsurveyed cities, we isolate the effects of neighborhood grading from broader urban trends. Our findings indicate that the maps reinforced racial sorting and segregation but did not independently drive the siting of environmental or climatic hazards. More broadly, the results illustrate the limits of correlational evidence in this setting: strong associations between historical grades and present-day environmental outcomes need not imply causal effects of the HOLC maps themselves. Distinguishing correlation from causation is therefore essential for understanding the legacy of redlining and for interpreting the role of historical housing policies in shaping contemporary environmental inequality.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder (2021) "The Effects of the 1930s HOLC "Redlining" Maps," *American Economic Journal: Economic Policy*, 13 (4), 355–392, 10.1257/pol.20190414.

Aaronson, Daniel, Daniel Hartley, Bhashkar Mazumder, and Martha Stinson (2023) "The long-run effects of the 1930s redlining maps on children," *Journal of Economic Literature*, 61, 846–862, 10.1257/jel.20221702.

Anderson, Michael L (2020) "As the Wind Blows: The Effects of Long-Term Exposure to Air Pollution on Mortality," *Journal of the European Economic Association*, 18 (4), 1886–1927, 10.1093/jeea/jvz051.

Bakkensen, Laura A. and Lala Ma (2020) "Sorting over flood risk and implications for policy reform," *Journal of Environmental Economics and Management*, 104, 102362, https://doi.org/10.1016/j.jeem.2020.102362.

Banzhaf, H Spencer and Randall P Walsh (2008) "Do people vote with their feet? An empirical test of tiebout," *American Economic Review*, 98, 843–863, 10.1257/aer.98.3.843.

Banzhaf, Spencer, Lala Ma, and Christopher Timmins (2019) "Environmental justice: The economics of race, place, and pollution," *Journal of Economic Perspectives*, 33, 185–208, 10.1257/jep.33.1.185.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Yongchen Zou (2024) "From Fog to Smog: The Value of Pollution Information," *American Economic Review*, 114 (5), 1338–81, 10.1257/aer.20200956.

Benz, Susanne Amelie and Jennifer Anne Burney (2021) "Widespread race and class disparities in surface urban heat extremes across the United States," *Earth's Future Research Article*, 9, e2021EF002016, 10.1029/2021EF002016.

Breiman, Leo (2001) "Random Forests," *Machine Learning*, 45 (1), 5–32, 10.1023/A:1010933404324.

Cain, Lucas, Danae Hernandez-Cortes, Christopher Timmins, and Paige Weber (2024) "Recent findings and methodologies in economics research in environmental justice," *Review of Environmental Economics and Policy*, 18, 116–142, 10.1086/728100.

California EPA (2021) "Pollution and prejudice: Redlining and environmental injustice in California," 8, https://storymaps.arcgis.com/stories/f167b251809c43778a2f9f040f43d2f5.

Chen, Shuai, Paulina Oliva, and Peng Zhang (2022) "The effect of air pollution on migration: Evidence from China," *Journal of Development Economics*, 156, 102833, 10.1016/j.jdeveco.2022.102833.

Christensen, Peter and Christopher Timmins (2022) "Sorting or steering: The effects of housing discrimination on neighborhood choice," *Journal of Political Economy*, 130, 2110–2163, 10.1086/720140.

Clay, Karen, Danae Hernandez-Cortes, Akshaya Jha, Joshua Lewis, Noah Miller, and Edson Severnini (2025) "The Social Lifecycle Impacts of Power Plant Siting in the Historical United States," technical report, National Bureau of Economic Research, 10.3386/w34109.

Currie, Janet, Lucas Davis, Michael Greenstone, and Walker Reed (2015) "Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings," *American Economic Review*, 105 (2), 678–709, 10.1257/aer.20121656.

Currie, Janet, John Voorheis, and Reed Walker (2023) "What caused racial disparities in particulate exposure to fall? New evidence from the Clean Air Act and satellite-based measures of air quality," *American Economic Review*, 113, 71–97, 10.1257/aer.20191957.

Cushing, Lara J., Shiwen Li, Benjamin B. Steiger, and Joan A. Casey (2022) "Historical redlining is associated with fossil fuel power plant siting and present-day inequalities in air pollutant emissions," *Nature Energy*, 8, 52–61, 10.1038/s41560-022-01162-y.

Depro, Brooks, Christopher Timmins, and Maggie O'Neil (2015) "White flight and coming to the nuisance: Can residential mobility explain environmental injustice?" *Journal of the Association of Environmental and Resource Economists*, 2, 439–468, 10.1086/682716.

Farber, Stephen (1998) "Undesirable facilities and property values: a summary of empirical studies," *Ecological Economics*, 24 (1), 1–14, https://doi.org/10.1016/S0921-8009(97)00038-4.

First Street Foundation (2022) "Property-level flood, fire and heat risk," *Dataset*.

Fishback, Price, Jessico LaVoice, Allison Shertzer, and Randall Walsh (2023) "The HOLC maps: How race and poverty influenced real estate professionals' evaluation of lending risk in the 1930s," *The Journal of Economic History*, 83, 1019–1056, 10.1017/S0022050723000475.

Hausman, Catherine and Samuel Stolper (2021) "Inequality, information failures, and air pollution," *Journal of Environmental Economics and Management*, 110, 102552, 10.1016/j.jeem.2021.102552.

Heblich, Stephan, Alex Trew, and Yanos Zylberberg (2021) "East-Side story: Historical pollution and persistent neighborhood sorting," *Journal of Political Economy*, 129, 1508–1552, 10.1086/713101.

Hollingsworth, Alex, Krzysztof Karbownik, Melissa A. Thomasson, and Anthony Wray (2024) "The Gift of a Lifetime: The Hospital, Modern Medicine, and Mortality," *American Economic Review*, 114 (7), 2201–38, 10.1257/aer.20230008.

Hsu, Angel, Glenn Sheriff, Tirthankar Chakraborty, and Diego Manya (2021) "Disproportionate exposure to urban heat island intensity across major US cities," *Nature Communications article*, 12, 2721, 10.1038/s41467-021-22799-5.

Hynsjö, Disa M and Luca Perdoni (2024) "Mapping out institutional discrimination: The economic effects of federal "redlining"," *CESifo Working Paper No. 11098*, 10.2139/ssrn.4820845.

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, 419–452, 10.1177/009614428000600404, doi: 10.1177/009614428000600404.

Jbaily, Abdulrahman, Xiaodan Zhou, Jie Liu, Ting-Hwan Lee, Leila Kamareddine, Stéphane Verguet, and Francesca Dominici (2022) "Air pollution exposure disparities across US population and income groups," *Nature*, 601, 228–233, 10.1038/s41586-021-04190-y.

Lane, Haley M, Rachel Morello-Frosch, Julian D Marshall, and Joshua S Apte (2022) "Historical redlining is associated with present-day air pollution disparities in U.S. cities," *Environmental Science and Technology Letters*, 9, 345–350, 10.1021/acs.estlett.1c01012.

Locke, Dexter H, Billy Hall, J Morgan Grove, Steward T A Pickett, Laura A Ogden, Carissa Aoki, Christopher G Boone, and Jarlath P M O'Neil-Dunne (2021) "Residential housing segregation and urban tree canopy in 37 US Cities," *npj Urban Sustainability*, 1, 15, 10.1038/s42949-021-00022-0.

Logan, John R. and Weiwei Zhang (2018) "Developing GIS maps for US cities in 1930 and 1940," in Ian Gregory, Don DeBats, and Don Lafreniere eds. *The Routledge Companion to Spatial History*, 1st edition, 229–249, Milton Park, Abingdon, Oxon; New York, NY: Routledge, 2018. Routledge, 10.4324/9781315099781-15.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles (2021) "National Historical Geographic Information System: Version 16.0," 10.18128/D050.V16.0.

McCormick, Erin (2022) "How Dividing US Cities Along Racial Lines Led to an Air Pollution Crisis 100 Years On," *The Guardian*, https://www.theguardian.com/us-news/2022/mar/09/redlining-air-pollution-us-cities, Newspaper article.

Michney, Todd M. (2022) "How the city survey's redlining maps were made: a closer look at HOLC's mortgage rehabilitation division," *Journal of Planning History*, 21 (4), 316–344, 10.1177/15385132211013361.

Michney, Todd M (2023) *How and why the home owners' loan corporation made its redlining maps*, Digital Scholarship Lab, https://dsl.richmond.edu/panorama/redlining/howandwhy, As found in Nelson (2023).

Miñano-Mañero, Alba (2024) "When are D-graded neighborhoods not degraded? Greening the legacy of redlining," 2, Mimeograph. Available at https://albaminanomanero.github.io/files/dgraded.pdf.

Moulton, Jeremy G., Nicholas J. Sanders, and Scott A. Wentland (2024) "Toxic Assets: How the Housing Market Responds to Environmental Information Shocks," *Land Economics*, 100 (1), 66–88, 10.3368/le.100.1.102122-0089R.

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 and Medicine*, 246, 112758, 10.1016/j.socscimed.2019.112758.

Nardone, Anthony, Joan A Casey, Rachel Morello-Frosch, Mahasin Mujahid, John R Balmes, and Neeta Thakur (2020) "Associations between historical residential redlining and current age-adjusted rates of emergency department visits due to asthma across eight cities in California: an ecological study," *The Lancet Planetary Health*, 4, e24–e31, 10.1016/S2542-5196(19)30241-4.

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, 17006, 10.1289/EHP7495.

Nelson, Robert K, LaDale Winling, Richard Marciano, and Nathan Connolly (2023) *Mapping Inequality: Redlining in New Deal America*, Digital Scholarship Lab, https://dsl.richmond.edu/panorama/redlining/, American Panorama: An Atlas of United States History.

Nembrini, Stefano, Inke R König, and Marvin N Wright (2018) "The revival of the Gini importance?" *Bioinformatics*, 34 (21), 3711–3718, 10.1093/bioinformatics/bty373.

Plumer, Brad and Nadja Popovich (2020) "How Decades of Racist Housing Policy Left Neighborhoods Sweltering," *The New York Times*, `https://www.nytimes.com/interactive/2020/08/24/climate/racism-redlining-cities-global-warming.html`, Interactive newspaper article; photographs by Brian Palmer.

Rothstein, Richard (2017) *The Color of Law A Forgotten History of How Our Government Segregated America*, Liveright.

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren (2024) "IPUMS Ancestry Full Count Data," `10.18128/D014.V4.0`, Dataset.

Ruggles, Steven, Nesile Ozder, Catherine A. Fitch et al. (2025) "IPUMS Multigenerational Longitudinal Panel," `10.18128/D016.V2.0`, Dataset.

Sager, Lutz and Gregor Singer (2025) "Clean identification? The effects of the Clean Air Act on air pollution, exposure disparities, and house prices," *American Economic Journal: Economic Policy*, 17, 1–36, `10.1257/pol.20220745`.

Schütt, Marvin (2021) "Systematic Variation in Waste Site Effects on Residential Property Values: A Meta-Regression Analysis and Benefit Transfer," *Environmental and Resource Economics*, 78 (3), 381–416, `10.1007/s10640-021-00536-2`.

Shen, Siyuan, Chi Li, Aaron van Donkelaar, Nathan Jacobs, Chenguang Wang, and Randall V Martin (2024) "Enhancing global estimation of fine particulate matter concentrations by including geophysical a priori information in deep learning," *ACS ES&T Air*, 1, 332–345, `10.1021/acsestair.3c00054`.

Shertzer, Allison, Tate Twinam, and Randall P Walsh (2022) "Zoning and segregation in urban economic history," *Regional Science and Urban Economics*, 94, 103652, `https://doi.org/10.1016/j.regsciurbeco.2021.103652`, Urban Economics and History.

Silva, Dakshina G De, Timothy P Hubbard, and Anita R Schiller (2016) "Entry and exit patterns of "toxic" firms," *American Journal of Agricultural Economics*, 98, 881–909, `10.1093/ajae/aaw012`.

Sood, Aradhya and Kevin Ehrman-Solberg (2023) "The long shadow of housing discrimination: Evidence from racial covenants," *SSRN Electronic Journal*, 1–52, `10.2139/ssrn.4606234`.

Tessum, Christopher W, Joshua S Apte, Andrew L Goodkind et al. (2019) "Inequity in consumption of goods and services adds to racial-ethnic disparities in air pollution exposure," *Proceedings of the National Academy of Sciences of the United States of America*, 116, 6001–6006, `10.1073/pnas.1818859116`.

Wing, Oliver E J, William Lehman, Paul D Bates et al. (2022) "Inequitable patterns of US flood risk in the Anthropocene," *Nature Climate Change*, 12, 156–162, `10.1038/s41558-021-01265-6`.

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, `10.1093/jahist/jaab066`.

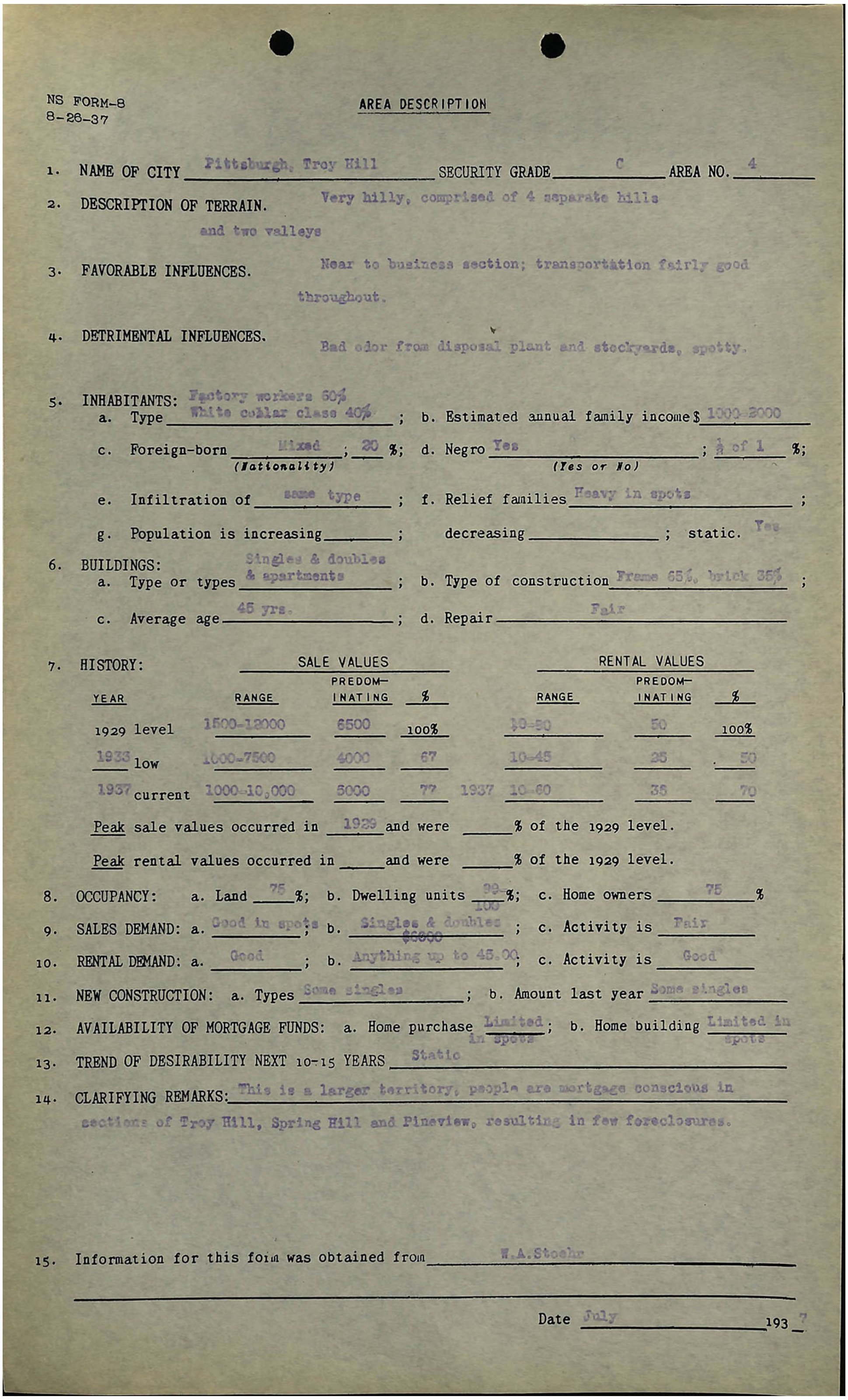
Wolverton, Ann (2009) "Effects of socio-economic and input-related factors on polluting plants' location decisions," *The B.E. Journal of Economic Analysis & Policy*, 9, `doi:10.2202/1935-1682.2083`.

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, 1036–1059, `10.1177/0096144211435126`, doi: 10.1177/0096144211435126.

Wright, Marvin N. and Andreas Ziegler (2017) "ranger : A Fast Implementation of Random Forests for High Dimensional Data in *C++* and *R*," *Journal of Statistical Software*, 77 (1), `10.18637/jss.v077.i01`.

A Appendix

Figure (A1) HOLC Area Description Form, Neighborhood C-4, Pittsburgh, PA



NS FORM-8
8-26-37

AREA DESCRIPTION

1. NAME OF CITY Pittsburgh, Troy Hill SECURITY GRADE C AREA NO. 4
2. DESCRIPTION OF TERRAIN. Very hilly, comprised of 4 separate hills and two valleys
3. FAVORABLE INFLUENCES. Near to business section; transportation fairly good throughout.
4. DETRIMENTAL INFLUENCES. Bad odor from disposal plant and stockyards, spotty.
5. INHABITANTS:
   a. Type Factory workers 60% White collar class 40%; b. Estimated annual family income $ 1000-2000
   c. Foreign-born Mixed (Nationality); 20 %; d. Negro Yes (Yes or No); ½ of 1 %;
   e. Infiltration of same type; f. Relief families Heavy in spots;
   g. Population is increasing ____; decreasing ____; static. Yes
6. BUILDINGS:
   a. Type or types Singles & doubles & apartments; b. Type of construction Frame 65%, brick 35%;
   c. Average age 45 yrs.; d. Repair Fair
7. HISTORY:

YEAR	SALE VALUES RANGE	SALE VALUES PREDOMINATING	SALE VALUES %		RENTAL VALUES RANGE	RENTAL VALUES PREDOMINATING	RENTAL VALUES %
1929 level	1500-12000	6500	100%		10-50	50	100%
1933 low	1000-7500	4000	67		10-45	25	50
1937 current	1000-10,000	5000	77	1937	10-60	35	70

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

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

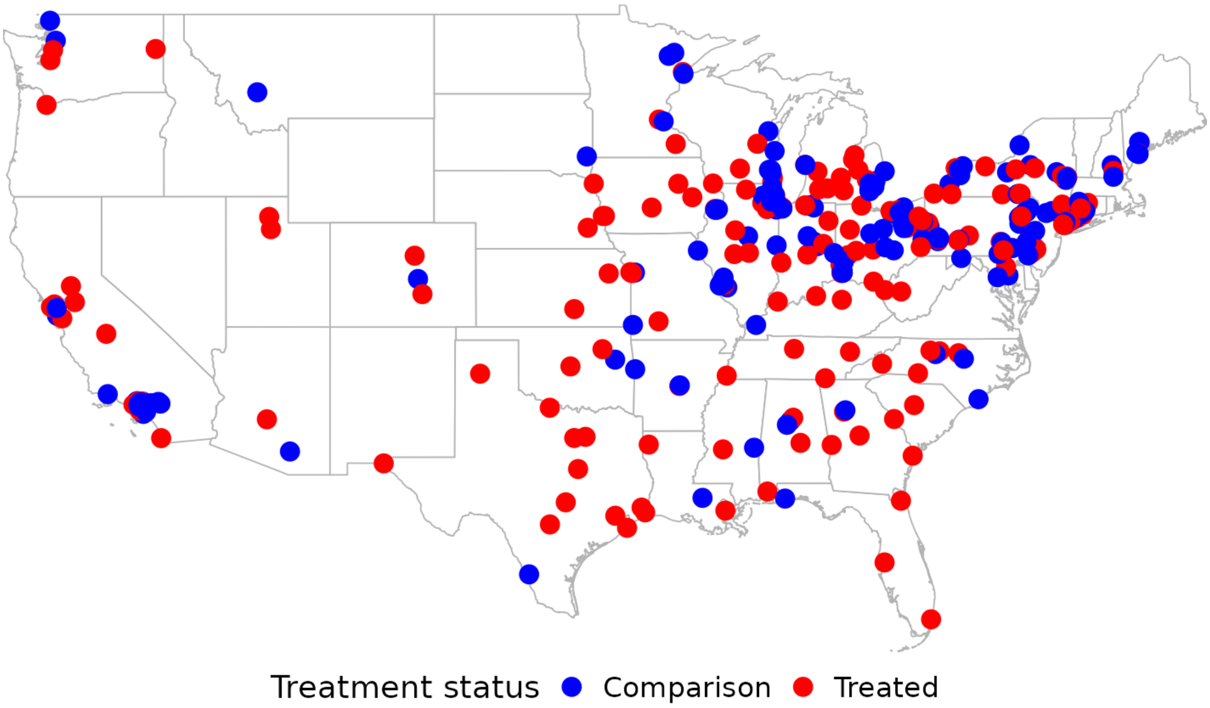
8. OCCUPANCY: a. Land 75 %; b. Dwelling units 99-100 %; c. Home owners 75 %
9. SALES DEMAND: a. Good in spots; b. Singles & doubles $6000; c. Activity is Fair
10. RENTAL DEMAND: a. Good; b. Anything up to 45.00; c. Activity is Good
11. NEW CONSTRUCTION: a. Types Some singles; b. Amount last year Some singles
12. AVAILABILITY OF MORTGAGE FUNDS: a. Home purchase Limited in spots; b. Home building Limited in spots
13. TREND OF DESIRABILITY NEXT 10-15 YEARS Static
14. CLARIFYING REMARKS: This is a larger territory, people are mortgage conscious in sections of Troy Hill, Spring Hill and Pineview, resulting in few foreclosures.

15. Information for this form was obtained from W.A.Stoehr

Date July 1937

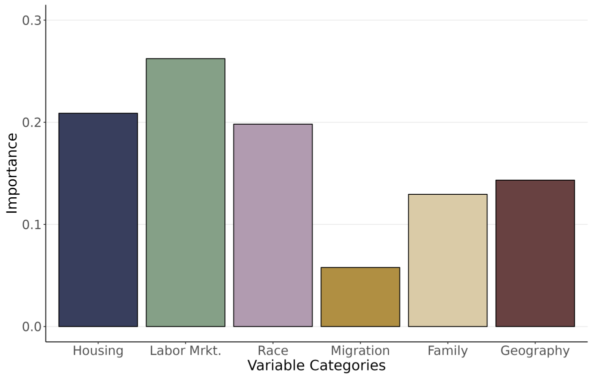
Notes: The scan of the *Area Description* for neighborhood C-4 of Pittsburgh, PA, has been provided by *Mapping Inequality* (?).

Figure (A2) Map of treated cities vs. control cities across the entire US

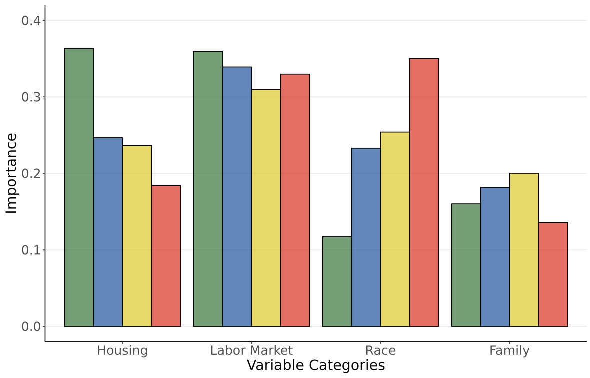


Notes: The map displays the geographic distribution of cities included in the analysis. Cities assigned to the control group are marked in blue, while cities in the treated group are indicated in red.

Figure (A3) Variable Importance Measures from Random Forest Algorithm



(a) For All HOLC Grades.



(b) Condition on HOLC Grades.

Notes: The histograms summarize the relative contribution of different groups of covariates to the random forest classification. The left panel displays category-level importance based on a bias-corrected impurity metric implemented in the `ranger` package (Wright and Ziegler, 2017) and discussed in Nembrini et al. (2018). The right panel reports permutation-based importance measures computed separately for each HOLC grade. For ease of comparison across panels, all importance measures are normalized.

Table (A1) Random Forest Performance, Confusion Matrix

		Data D	Data C	Data B	Data A
Prediction	D	12923	651	85	12
	C	967	18056	920	101
	B	86	581	8405	386
	A	8	26	123	2097
Accuracy		91.31%			
Class Sensitivity		92.41	93.49	88.17	80.77
Prevalence		30.78	42.52	20.99	5.71
Detection Prev.		30.09	44.12	20.82	4.96

Notes: The table reports a comparison between observed HOLC grades and grades predicted by the random forest classifier using an out-of-sample test set. The test sample consists of 25% of the original hexagon-level observations and is drawn using stratified random sampling by city population and observed HOLC grade. The unit of observation is a hexagon, and the sample includes all hexagons located in mapped cities with at least 20 residents in 1930. Predicted grades correspond to the modal class assigned by the trained random forest model. *Overall accuracy* denotes the share of hexagons for which predicted and observed grades coincide. *Class sensitivity* for grade j is defined as the fraction of correctly classified hexagons among those with observed grade j. *Prevalence* reports the distribution of observed grades in the test sample, while *detection prevalence* reports the distribution of predicted grades.

Table (A2) Balance of neighborhood characteristics on C-B

Sample:	1930			
	Black share (1)	Ownership rate (2)	House value (3)	Rent price (4)
LowGrade × Treated	0.01	0.13	0.06	0.03
	(0.03)	(0.11)	(0.08)	(0.10)
LowGrade	0.06***	-0.64***	-0.59***	-0.19*
	(0.02)	(0.06)	(0.05)	(0.10)
Dependent variable mean	-0.22	0.16	0.06	0.05
Observations	15,965	15,957	15,927	15,927
City fixed effects	✓	✓	✓	✓

Sample:	1920		1910	
	Black share (1)	Ownership rate (2)	Black share (3)	Ownership rate (4)
LowGrade × Treated	0.01	0.09	0.04	-0.01
	(0.03)	(0.09)	(0.06)	(0.09)
LowGrade	0.03*	-0.53***	-0.03	-0.38***
	(0.01)	(0.05)	(0.03)	(0.07)
Dependent variable mean	-0.20	0.18	-0.17	0.15
Observations	12,611	12,607	10,505	10,499
City fixed effects	✓	✓	✓	✓

Notes: Regression of neighborhood characteristics in 1930, prior to redlining. The observations are weighted by 1930 neighborhood population. Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01

Table (A3) African American segregation: neighborhoods

Sample:	D-C		C-B	
	1940	2010	1940	2010
	(1)	(2)	(3)	(4)
LowGrade × Treated	0.012	0.017	0.004	0.029*
	(0.037)	(0.022)	(0.006)	(0.017)
LowGrade	0.109***	0.069***	0.010***	0.032**
	(0.024)	(0.015)	(0.003)	(0.013)
Dependent variable mean	0.07	0.23	0.02	0.17
Observations	17,984	30,836	15,394	26,146
City fixed effects	✓	✓	✓	✓

Notes: Estimation of Share of African American inhabitants of a neighborhood. Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01

Table (A4) Rent: neighborhoods

Sample:	D-C		C-B	
	1940	1990	1940	1990
	(1)	(2)	(3)	(4)
LowGrade × Treated	0.126	-2.678	12.955	1.105
	(12.423)	(5.021)	(12.895)	(5.119)
LowGrade	-15.197	-17.648***	2.114	-28.433***
	(9.315)	(3.237)	(6.654)	(3.474)
Dependent variable mean	38.16	340.30	43.14	359.35
Observations	17,970	30,791	15,381	26,117
City fixed effects	✓	✓	✓	✓

Notes: Estimation of average rent in a neighborhood. Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01

Table (A5) Home ownership: neighborhoods

Sample:	D-C		C-B	
	1940	2010	1940	2010
	(1)	(2)	(3)	(4)
LowGrade × Treated	0.003	-0.020	0.030	0.013
	(0.023)	(0.020)	(0.025)	(0.025)
LowGrade	-0.067***	-0.055***	-0.131***	-0.084***
	(0.014)	(0.014)	(0.017)	(0.019)
Dependent variable mean	0.45	0.46	0.51	0.50
Observations	17,981	30,832	15,391	26,142
City fixed effects	✓	✓	✓	✓

Notes: Estimation of home ownership rate in a neighborhood. Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01

Table (A6) African American segregation: neighborhoods, long-differences

Sample:	D-C		C-B	
	1930-1940	1930-2010	1930-1940	1930-2010
	(1)	(2)	(3)	(4)
LowGrade × Treated	0.012**	0.001	0.001	0.059***
	(0.005)	(0.031)	(0.002)	(0.022)
LowGrade	-0.007**	-0.034**	0.000	0.021
	(0.003)	(0.015)	(0.001)	(0.014)
Dependent variable mean	0.00	0.14	0.00	0.13
Observations	17,916	18,661	15,340	15,965
City fixed effects	✓	✓	✓	✓

Notes: Estimation of Share of African American inhabitants of a neighborhood. Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01

Table (A7) Air pollution and climatic hazards, C vs. B

Sample:	C-B			
Year:	2020			
	Flood	Heat	NO_2	$PM_{2.5}$
	(1)	(2)	(3)	(4)
LowGrade × Treated	0.17	-0.06*	0.09	-0.03
	(0.13)	(0.03)	(0.21)	(0.03)
LowGrade	0.17*	0.09***	-0.07	0.04*
	(0.09)	(0.03)	(0.16)	(0.02)
Dependent variable mean	1.50	4.28	39.54	7.64
Observations	25,716	25,716	26,156	26,156
City fixed effects	✓	✓	✓	✓

Notes: Standard errors are clustered at the city level. * 0.1 ** 0.05 *** 0.01