The Lingering Legacy of Redlining on School Funding, Diversity, and Performance

Christopher Cleveland[1]
Harvard University
chcleveland@g.harvard.edu

Dylan Lukes
Harvard University
dylanlukes@g.harvard.edu

October 2022

Abstract

Between 1935-1940 the Home Owners' Loan Corporation (HOLC) assigned A (minimal risk) to D (hazardous) grades to neighborhoods that reflected their lending risk from previously issued loans and visualized these grades on color-coded maps. These maps arguably influenced mortgage lenders to provide or deny home loans within residential neighborhoods. In this study, we leverage a spatial analysis of 144 HOLC-graded core-based statistical areas (CBSAs) to understand how HOLC maps relate to current patterns of district and school funding, school racial diversity, and school performance. We find that districts composed of D neighborhoods have less district per-pupil total revenues, but schools in D neighborhoods have higher per-pupil total expenditures. We find that schools in D neighborhoods have larger shares of Black and non-White student bodies, larger shares of low-income students, and worse average test scores. We also document persistence in these patterns across time. These findings suggest that policymakers need to consider the historical implications of redlining and neighborhood inequality on neighborhoods today when designing modern interventions focused on improving the outcomes of students of color and students from low-income backgrounds.

[1] Corresponding author (chcleveland@g.harvard.edu).

1 Introduction

The United States has a long history of racially discriminatory policies and practices. Redlining, where the Home Owners' Loan Corporation (HOLC) assigned A-D security ratings to nearly 240 cities across the United States between 1935 and 1940, is a small but important part of this history.[2] With the advent of recently digitized HOLC maps by the "Mapping Inequality Project" headed by the University of Richmond's Digital Scholarship Lab, the long-run impact of HOLC A-D security ratings on social and economic outcomes has become increasingly well documented by social scientists. While much of the literature today shows HOLC maps' adverse effects on modern outcomes such as credit scores, family structure, homeownership, home values, household income, neighborhood segregation, and incarceration (Appel & Nickerson, 2016; Krimmel, 2018; Anders, 2018; Aaronson et al. 2020; Aaronson et al., 2021), very few studies, if any, look at the long-term relationship between HOLC maps and contemporary educational outcomes. We hypothesize that the association between educational outcomes and neighborhood quality is intertemporal such that U.S. neighborhood inequality from the early 20th century, often resultant from historical discriminatory practices, is predictive of present-day educational outcomes. With recent work by Aaronson et al. (2020) and Aaronson et al. (2021) establishing a causal relationship between HOLC maps and neighborhood segregation, housing, and socioeconomic outcomes, it would be surprising if these results did not spill over to measures of educational quality given the deep-rooted connection between neighborhoods and the schools and districts that serve them (Massey & Denton, 1993; Aaronson, 1998; Chetty et al.,

[2] HOLC and Federal Housing Administration (FHA) maps are not the same. HOLC maps were created first in 1935 under the Federal Home Loan Bank Board (FHLBB) as part of a systematic appraisal process by the HOLC to evaluate the lending risk for loans they had already issued to mitigate non-farm mortgage defaults and foreclosures during the Great Depression (Fishback et al., 2020; Aaronson et al., 2020). The FHA, under the U.S. Department of Housing and Urban Development, created its maps and used them to determine whether homes qualified for mortgage insurance. While historians agree that the HOLC maps influenced the FHA maps, it is unknown to what degree since all but two FHA maps (Chicago, IL; Greensboro, NC) have been lost or destroyed (Light, 2010).

2016; Bayer et al., 2020; Dalane and Marcotte, 2020). To this end, our paper addresses the following questions: (1) How do historic HOLC D neighborhoods relate to existing patterns of district-level and school-level funding, student racial diversity, and student performance? (2) Do these patterns vary by region? And (3) How, if at all, do these patterns vary over time? In doing so, we hope to contribute to the existing body of research connecting HOLC maps to modern-day outcomes by providing some of the first evidence of the long-term association between 1935-1940 HOLC A-D grades and educational outcomes.

To link historical HOLC maps to contemporary educational institutions, we map 1935-1940 HOLC A-D neighborhood grades to present-day schools and districts. For schools, individual HOLC A-D grades are assigned based on overlapping school-level latitude and longitude geolocations and HOLC A-D geospatial polygons. For districts, we determine HOLC A-D mappings using the area, in square miles, of HOLC A-D polygons that overlap with each respective district boundary. We use these mappings to analyze the relationship between HOLC A-D security ratings and our district and school outcomes, both modern-day and over time, using various plots and statistical tests.

Our analysis identifies important findings. First, for district-level funding, present-day districts located in D neighborhoods have less district-level per-pupil total revenues but higher per-pupil federal and state revenues than those mapped to HOLC A, B, and C security ratings. However, the differences in per-pupil federal and state revenues are not large enough to overcome the sizeable gaps in per-pupil local funding that favor those districts mapped to higher HOLC security ratings. These local, state, and federal dynamics drive the differences in per-pupil total revenue by HOLC A-D grades. Conversely, at the school level, we find that schools mapped to HOLC D ratings have, on average, more per-pupil total expenditures and per-pupil federal expenditures than schools mapped to HOLC A, B, and C security ratings. Differences in aggregated per-pupil state and local expenditures between D vs. A, B, and C schools are statistically insignificant and signal the countervailing force that state funding has on local funding, which favors higher-rated HOLC security ratings. In reconciling the district and school finance results, we find that schools mapped to HOLC D grades have the largest shares of students qualifying for free and reduced-price lunch, making them eligible for Title I funding, the largest federal funding program for U.S. public schools. This finding provides a plausible mechanism for why HOLC D schools, regardless of what HOLC A-D district they fall within, have the most per-pupil funding of all HOLC A-D mapped schools. Furthermore, our district-level results remain pertinent as they represent an upper limit on how much a district can allocate to their most-in-need schools (i.e., D schools) while still maintaining funding levels for their other district schools.

Second, for student racial diversity, we find that modern-day schools located in HOLC D neighborhoods have larger school-level shares of Black and non-White student bodies and less diverse student populations than their more highly rated HOLC A, B, and C counterparts, nationally and by region. While schools in higher-rated HOLC A, B, and C neighborhoods have larger shares of White student bodies, they also have, on average, more diverse student populations. Even more, school diversity is monotonically increasing in HOLC rating, such that A schools have the greatest levels of diversity, followed by B, C, and ending with D.

Third, for student performance, we find that schools today located in historic HOLC D areas have worse school-level average math and reading scores than their more highly rated A, B, and C peers, nationally and by region. We also observe minor differences in either measure of average learning rates or educational opportunity changes by A-D HOLC grade. These results are true nationwide and by region. Finally, we document positive time trends for the finance and diversity outcomes across all HOLC A-D grades from the late 1980s to today, but persistent and widening gaps between schools in historically redlined D neighborhoods and those in A, B, and C neighborhoods. We do not document the time series trends for our school performance outcomes since measures are only available as pooled estimates spanning 2009 to 2018.

In summary, our paper sheds light on the previously unexplored relationship between 1935-1940 HOLC maps and modern district and school-level outcomes. Overall, our paper provides some of the first evidence documenting the association between historic HOLC A-D grades, which captured neighborhood inequality in the 1930s and subsequently contributed to neighborhood inequality in the following generations, and modern-day educational outcomes. While we cannot definitively say whether HOLC redlining caused the modern-day educational inequalities we show in this paper, HOLC A-D grades' predictive ability hint at a stubborn historical legacy. We believe these results indicate the need for educational policymakers to consider the historical implications of past neighborhood inequality on present-day neighborhoods when designing modern education interventions focused on improving the life outcomes of students of color and students from socioeconomically disadvantaged communities.

The remainder of this paper is as follows. Section 2 provides an overview of the history of HOLC maps, the current set of literature that links HOLC security ratings to social and economic outcomes, and related literature on neighborhoods and schools. Section 3 describes the data used for this research. Section 4 and Section 5 outline this paper's methods, including analysis samples construction, and details the analytic approach. Section 6 provides the results and a discussion. Section 7 concludes.

2 Literature

2.1 HOLC Maps History

In 1932, the Federal Home Loan Bank Board (FHLBB) was created to manage federal savings and loan associations. In 1933, under the FHLBB, the Home Owners' Loan Corporation (HOLC) agency was subsequently made to oversee the troubled U.S. mortgage market and tasked with purchasing and refinancing non-farm mortgages to limit foreclosures and defaults. After the HOLC completed issuing loans to distressed properties, to evaluate their lending risk, they created a systematic appraisal process that included neighborhood-level characteristics such as race, ethnicity, immigration status, household income, homeownership rates, access to public services, and occupation type (Hillier, 2003; Crossney & Bartelt, 2005; Fishback et al., 2020). Between 1935 and 1940, HOLC's department of Research and Statistics used thousands of realtors, developers, lenders, and appraisers to create neighborhood-by-neighborhood security ratings of 239 cities and made over five million appraisals (Hillier, 2003; Crossney & Bartelt, 2005). Neighborhoods were graded on a scale of A (i.e., least risky/most stable) to D (i.e., most risky/least stable) based upon the perceived risk of making housing loans in different neighborhoods. These grades were later solidified as color-coded maps in which D-graded areas were colored red and prompted the later coinage of the term "redlining."

Each neighborhood had a standardized assessment sheet used to assign the HOLC grades. For example, area descriptions used in Los Angeles County in 1939 included eight sections. 1) *Population* asked for a record of whether the population was increasing, decreasing, or static. It also asked for the class and occupation of residents, the percentage of foreign families and their nationalities, the percentage of Black families, and whether the population trends reflected shifting or infiltration. 2) *Buildings* asked for the type and size of the building, construction, average age, repair status, occupancy rate, owner-occupied, 1935 price bracket, 1937 price bracket, 1939 price bracket, sales demand, predicted price trend, 1935 rent bracket, 1937 rent bracket, 1939 rent bracket, rental demand, and predicted rent trend. 3) *New Construction* recorded the number of new properties built within the past year, the prices for these units, and how they were selling. 4) *Overhang of Home Properties* captured unsold HOLC properties. 5) *Sale of Home Properties* captured sold HOLC properties 6) *Mortgage Funds* captured mortgage funds' availability. 7) *Total Tax Rate per $1000* captured the local tax rate. 8) *Description and Characteristics of Area* captured other qualitative detail about the terrain and population.[3]

A current debate exists as to whether HOLC A-D grade assignments were racially biased or merely a geographic snapshot of an outcome caused by America's long history of racial discrimination before it (Fishback et al., 2020). There is evidence that race, immigration status, household income, and ethnicity were often explicit factors in the HOLC grading process, such that non-White racial groups, immigrants, low-income households, and ethnic minorities were more likely to receive lower grades (Jackson, 1980; Connolly, 2014; Nelson et al., 2020). However, the degree to which HOLC maps influenced the lending practices and underwriting standards of the FHA and other private lenders is an ongoing debate. Some authors argue that access to HOLC maps was limited, while others argue that access to HOLC maps was ubiquitous and materially influenced public and private lending policies and practices (Jackson, 1980; Hillier, 2003; Light, 2010; Woods, 2012). Regardless, it is clear that HOLC encouraged the general rule of using maps to classify the creditworthiness and lending risks of neighborhoods and broadly are considered to have redirected public and private capital and homeownership for intergenerational wealth-building to native-born white families and away from African American and immigrant families (Appel & Nickerson, 2016; Krimmel, 2018; Anders, 2018; Aaronson et

[3] The area descriptions for Los Angeles A-1 and D-1 can be viewed in Figure A.1 in Appendix A.

al., 2020; Aaronson et al., 2021).

2.2 Modern Outcomes Linked to HOLC Maps

A literature base that leverages both descriptive and causal empirical strategies is developing that links neighborhoods' HOLC grades to various contemporary outcomes. Mitchell and Franco (2018) descriptively analyze HOLC-graded areas' modern demographic and residential patterns. They find that many neighborhoods rated high-risk or “Hazardous” by HOLC eight decades ago are low-to-moderate income (~74%) and minority neighborhoods (~64%) today. Additionally, the authors find greater economic inequality, higher levels of interaction between Black and White residents, and a stronger positive association between gentrification and economic change in neighborhoods rated “Hazardous” by HOLC in their sample. Regionally, the South showed the smallest change in the HOLC-rated “Hazardous” neighborhoods with lower incomes and more majority-minority residents. In addition, the Midwest region was found to closely mirror the South in the persistence of low-to-moderate income neighborhoods in HOLC “Hazardous” areas.

Three papers have highlighted the impacts of the HOLC maps on housing and socioeconomic outcomes. Appel and Nickerson (2016) use a spatial regression discontinuity boundary design to study long-term impacts on home prices. They show that housing characteristics varied smoothly at the boundaries when the maps were created. Despite this initial smoothness, they find that HOLC "redlined" neighborhoods have about 4.8% lower home prices, fewer owner-occupied homes, and more vacant buildings relative to adjacent areas. Similarly, Aaronson, Hartley, and Mazumder (2020) use a spatial RD boundary propensity score design to study the effects of the HOLC A-D maps on the long-run trajectories of neighborhoods. They find that the maps led to reduced homeownership rates, house values, and rents and increased racial segregation in later decades. They conclude that the HOLC maps had sizeable and persistent impacts on urban-neighborhood development driven by reduced credit access and the subsequent neighborhood disinvestment (Aaronson et al., 2020). In a follow-up paper, Aaronson, Faber, Hartley, Mazumder, and Sharkey (2021) extend Aaronson et al. (2020) to study the impact of HOLC A-D maps on a variety of socioeconomic outcomes such as income rank, family structure, incarceration, and geographic mobility. Using data from the Opportunity Atlas and an identification strategy like Aaronson et al. (2020), they find sizeable and statistically significant differences across all outcomes that favor higher-rated HOLC neighborhoods over lower-rated ones.

Other papers have explored a broader set of outcomes connected to the HOLC maps. Jacoby, Dong, Beard, Wiebe, and Morrison (2018) use descriptive spatial analysis to evaluate the relationship between HOLC grades and modern firearm violence in Philadelphia, PA. Using data from the 1940 U.S. Census the authors adjust for socio-demographic factors at the time of HOLC-map creation and find that firearm injury rates are highest in historically HOLC D areas of Philadelphia (Beard et al., 2017; Jacoby et al., 2018). Nardone, Casey, Rudolph, Karasek, Mujahid, and Morello-Frosch (2020) descriptively explore how birth outcomes within California vary based upon HOLC grade. The authors find that worse HOLC ratings are associated with adverse birth outcomes. However, their findings are inconsistent when using propensity score matching and stratifying by metropolitan area (Nardone et al., 2020). Hoffman, Shandas, and Pendleton (2020) use descriptive spatial analysis to study the association between a neighborhood's HOLC grade and modern-day surface temperatures. They find that for most areas in their study, previous HOLC redlined areas have consistently elevated land surface temperatures compared to non-redlined areas. Cities in the Southeast and Western regions display the largest differences in land surface temperatures, while those in the Midwest region show the least. Overall, the authors find that nationally, land surface temperatures in previously HOLC redlined locations are about 2.6 °C hotter than those in non-redlined ones.

These recent studies suggest a strong association between historical practices for assigning grades to neighborhoods and modern outcomes. The HOLC maps and related historical policies have contributed to racial disparities in neighborhood diversity, income, health care, access to healthy food, incarceration, and public infrastructure investment. This paper's main contribution is to explore the relationship between HOLC map grades and present-day educational outcomes of school and district funding, school racial diversity, and school performance. As discussed in the following section, these outcomes are intricately connected to U.S. housing and neighborhoods and have yet to be explored.

2.3 The Relationship Between Schools, Neighborhoods, and Student Outcomes

The academic literature highlights the inter-relatedness of school funding, school racial diversity, and broader neighborhood contexts on students' long-term outcomes. Education funding in the U.S. operates at federal, state, local, and district levels. The two primary federal revenue sources are Title I and IDEA, which distribute dollars based on student population size and poverty concentration. At the state level, each state has different mixes of revenue streams dedicated to education and implements a unique funding formula to direct dollars to school districts. At the local level, districts primarily use property taxes to generate revenue. Within a district, funding is often distributed through a traditional centralized model in which the district deploys resources to schools in the form of staff, programs, and services (Urban Institute, 2017; Roza, Hagan & Anderson, 2020; Baker et al., 2018; Brittain, Willis & Cookson, 2019; Green et al., 2021).

A recent study has illuminated the connection between school quality and neighborhood value. Using multiple decades of U.S. housing price data, Bayer et al. (2020) conduct a national study on the causal effect of school spending and local taxes on housing prices. The authors find that households value school spending, including spending on teachers and staff salaries. Additionally, they show that salary spending is allocated inefficiently throughout the U.S. and find that within-district-salary-expenditure raises funded via local taxes would increase home prices.

Interrelatedly, there has been a growing focus on race- and income-based segregation within districts and schools. Dalane and Marcotte (2020) find that classroom segregation increased by around 10 percent between 2007 and 2014 for those elementary and middle schools in their study. Further, they find that segregation of low-income students within schools is associated with the level of segregation between schools in districts and that this correlation grew stronger across the panel years. These findings of increasing segregation in certain parts of the country have been highlighted in other recent studies (Clotfelter et al., 2019; Alcaino & Jennings, 2020; Clotfelter et al., 2021; Monarrez et al., 2020). The salience of neighborhoods became well documented through the efforts of Massey & Denton (1993) to link persistent poverty among Black people in the United States to the unparalleled degree of deliberate segregation they experience in American cities. Chetty, Hendren, and Katz (2016) found from the Moving to Opportunity experiment that a reduction in neighborhood poverty had no impact on short-term reading and math test scores for Black children but improved early adult outcomes, including educational attainments and labor market earnings. More broadly, the recent development of the Opportunity Atlas shed light on the average adult outcomes of individuals who grew up in each U.S. Census tract to trace back the roots of poverty and incarceration

(Chetty et al., 2018; Chetty & Hendren, 2018a; Chetty & Hendren, 2018b; Murnane, 2021).

Given this literature base, this paper's primary contribution is to extend the roots of perceived issues in school funding, racial diversity, and performance to the inception of HOLC redlining to demonstrate the long-term relationship between HOLC A-D security ratings and present-day educational outcomes.

3 Data

In what follows, we provide a succinct overview of the Home Owners' Loan Corporation (HOLC) maps created by the "Mapping Inequality" project (Nelson et al., 2020), and of the district and school-level data sources we used in our analysis. For further details, please refer to Appendix Table A.1-A.3.

3.1 HOLC Data

The 1935-1940 HOLC maps were preserved by the U.S. National Archives and recently digitized by a team at the University of Richmond, Virginia Tech, University of Maryland, and Johns Hopkins University as part of the "Mapping Inequality" project (Nelson et al., 2020). This data contains digitized versions of every available city-level HOLC map from 1935 to 1940 in Shapefile or GeoJSON format. The spatial data includes over 7,000 neighborhoods and 240 unique cities across the United States. For each HOLC city, the data contains HOLC Geographic Information System (GIS) polygons, HOLC grade assignment text (i.e., A, B, C, D), and detailed area description transcriptions (Nelson et al., 2020). We also leverage digitized area description information about the HOLC polygons (Markley, 2021). For our study, we aggregate the data to 144 unique core-based statistical areas (CBSAs), which span the United States and include all four U.S. Census Bureau Regions (i.e., Midwest, Northeast, South, West). These CBSAs are captured below in Figure 1. Finally, we include all CBSAs with at least one HOLC grade in our analysis sample to maximize geographic coverage. As a result, some CBSAs do not contain a full A-D HOLC rating set.

[Insert Figure 1 here]

3.2 District and School Data

In this paper, we use geospatial and non-geospatial school and district-level data. Our school and district-level geospatial data consists of 2018-2019 district boundary data derived from the 2019 Census TIGER/Line Shapefiles and school-point location data from the 2018-19 NCES EDGE program (U.S. Census Bureau, 2020).

This paper's non-geospatial data includes measures of school and district level finances, demographics, and academic performance. For the district-level fiscal outcomes, we leverage the most recent NCES 2017-18 F-33 survey data, which provides general financing information (e.g., revenue and expenditure totals and subtotals) at the district level (U.S. Department of Education, 2020). For school-level finance data we use the newly released National Education Resource Database on Schools (NERD$) data includes state-reported 2018-19 school-level expenditures across 49 states and the District of Columbia and captures per-pupil total expenditures, federal expenditures, and combined state and local expenditures.

For school-level demographics and additional non-fiscal measures captured by NCES, we use NCES 2017-2018 and NCES 2018-19 data. Finally, for our school performance analyses, we use school-level data from the Stanford Education Data Archive (SEDA). This data provides students' academic outcomes in grades 3-8, spanning SY2008-09 to SY2017-18, and includes students' average test scores, test score trends, and learning rates as defined by SEDA.[4]

4 Analysis Samples

[4] See the data appendix for more information on each data source.

We use an approach motivated by Hoffman, Shandas, and Pendleton (2020), which documents the association of HOLC redlining policies on resident exposure to intra-urban heat. While our papers' methods deviate, the overarching analytic templates are similar.[5] In what follows, we provide an overview of how we constructed each school and district-level analysis sample and describe our analytic approach.

[Insert Table 1 here]

4.1 School-Level Analysis Samples

We connect the geospatial HOLC maps and the geospatial NCES public district and school-level data. If a modern-day school is contained within a historic HOLC A-D polygon based on its latitude and longitude point location, it is assigned that historic HOLC A-D polygon grade. If not, no grade is assigned, and the school is dropped from the sample.[6] We develop a slightly distinctive sample for each outcome based upon available school-level data. For a more thorough overview of the school-level analysis samples, please refer to Appendix Table A.1 Panel A and Panel B.

4.2 District-Level Analysis Sample

For the NCES district-level geospatial data, our approach is more nuanced. Unlike school point locations, district boundaries often span large metropolitan areas or entire counties, and as a result, at times envelope multiple HOLC A-D polygons. We use a one-to-many mapping procedure for districts. We first overlay all 2018-19 U.S. public school district boundaries with 1935-1940 HOLC A-D maps. If a HOLC A-D polygon overlaps, even in part, with a district

[5] Hoffman, Shandas, and Pendleton (2020) map intra-urban land surface temperature anomalies to HOLC ratings, whereas this paper maps school and district outcomes to HOLC ratings. In addition, instead of calculating a "delta" variable that consists of demeaned HOLC A-D averages within a given city by their respective city-wide average, we use CBSA fixed effects.

[6] Specifically, we execute a "one to many" join via the ArcGIS "completely contains" spatial relationship option. This option matches features from disparate data only if a target layer features from one dataset completely contain join-layer features from another. HOLC A-D ratings and public primary and secondary schools are the target and join layers for this exercise, respectively.

boundary, we link it to that overlapping district and calculate the intersecting area of the HOLC A-D polygon. Any district that does not intersect with at least one HOLC A-D polygon is dropped from the sample. Once this mapping exercise is complete, we have a dataset consisting of (1) U.S. public school districts, (2) HOLC A-D polygons that overlapped, entirely or in part, with said U.S. public school districts, and (3) HOLC A-D polygon areas, in square miles, based on the intersecting area between a HOLC A-D polygon and a district boundary.

We use this data to calculate the HOLC A-D weighted average for each public school district, where weights are derived from the overlapping areas (in square miles) of the HOLC polygons and the district boundaries. The weighted average formula can be viewed below:

$$\overline{HOLC_{LEA}} = \frac{1 * \sum_a^A Area_{a \cap LEA} + 2 * \sum_b^B Area_{b \cap LEA} + 3 * \sum_c^C Area_{c \cap LEA} + 4 * \sum_d^D Area_{d \cap LEA}}{\sum_a^A Area_{a \cap LEA} + \sum_b^B Area_{b \cap LEA} + \sum_c^C Area_{c \cap LEA} + \sum_d^D Area_{d \cap LEA}}$$

As expected, for each district, the weights taken together sum to one, where (1) the denominator is the total sum of all overlapping HOLC-district areas in a district and (2) the numerator is the sum of the overlapping HOLC-district areas for each HOLC A-D grade multiplied by the numeric of that HOLC grade. This procedure produces a continuous set ranging from one to four that we then discretize and link directly to a singular HOLC A-D grade which we use later in our analysis of district-level finances.[7]

We merge this dataset with the 2017-18 NCES district-level F-33 fiscal data and removing districts with missing F-33 finance data, we get our final cross-sectional "Fiscal" district-level analysis sample. We provide a more thorough overview of analysis samples in the Appendix A (Table A.1 Panel C).[8]

[7] We considered several potential weighting strategies and sample inclusion strategies that did not materially affect the findings. This strategy aligns with the methods employed by other papers leveraging HOLC data.

[8] In addition, to view examples of HOLC A-D maps overlaid with NCES district boundaries, please refer to Figure A.2 in Appendix A, where we provide NCES district boundary maps superimposed on a 1935-1940 HOLC map for Los Angeles Unified School District.

5 Empirical Strategy for Outcomes

Across our outcomes, our primary empirical strategy is inspired by a difference of means approach taken by Hoffman, Shandas, and Pendleton (2020), which documents the association of HOLC redlining policies with resident exposure to intra-urban heat.[9] While similar, we base our approach on a CBSA fixed effects regression model with HOLC grade indicators, where we are interested in determining whether, if at all, school and district outcomes for those located in once redlined HOLC D neighborhoods differ from those found in historically HOLC A, B, or C neighborhoods.[10] This approach allows us to quantify, for a given urban region, how much better or worse off schools or districts with a particular HOLC grade assignment (e.g., D) are from the set of all other HOLC schools or districts (e.g., A, B, C) in that urban region. Given the heterogeneity in state funding mechanisms across the U.S. (which have implications for school and district funding levels and distributions locally) and nationwide differences in demographic compositions, a within CBSA comparison of our outcomes by HOLC A-D grades is preferred to an unadjusted one. Our use of CBSA fixed effects also ensures that differences in any time-invariant covariates at the city-level are differenced out in our model, which we present below:

$$Y_{ij} = \alpha + \beta_1 HOLC_{A_{ij}} + \beta_2 HOLC_{B_{ij}} + \beta_3 HOLC_{C_{ij}} + \gamma_j + \epsilon_i$$

where i indexes schools or districts, j indexes CBSAs, and γ_j is the CBSA fixed effect.[11] For schools, $HOLC_A$, $HOLC_B$, and $HOLC_C$ are the estimated average differences in the outcome Y_{ij} between schools in what once were HOLC D graded neighborhoods and their contemporaries

[9] Hoffman, Shandas, and Pendleton (2020) map intra-urban land surface temperature anomalies to HOLC ratings, whereas this paper maps school and district outcomes to HOLC ratings. In addition, instead of calculating a "delta" variable that consists of demeaned HOLC A-D averages within a given city by their respective city-wide average, we use CBSA fixed effects.

[10] Recall, for districts, this is not the case. Districts are assigned a composite HOLC A-D grade based on the weighted average of HOLC A-D polygons they contain, where weights are based on the area, in square miles, of overlapping district boundaries and HOLC A-D polygons. See the above section for details. We explored alternative fixed effects specifications including state fixed effects or no fixed effects but are comfortable within the "within-city" interpretation provided by the CBSA fixed effects strategy.

located in what once were HOLC A, B, or C HOLC graded neighborhoods, respectively. For districts, the interpretation is similar but with the caveat that HOLC A-D grades represent HOLC A-D weighted averages, which could but often do not refer to a single HOLC A-D polygon, but instead a collection of them.

The key identifying assumption of our approach is that within-CBSA selection into a HOLC A-D grade assignment for schools or districts is uncorrelated with (1) uncontrolled for time-variant determinants of our outcomes and (2) unobservable determinants of our outcomes. Threats of validity arise if $E\left(\epsilon_i \middle| HOLC_A, HOLC_B, HOLC_C, \gamma_j\right) \neq 0$, which is true in the presence of school or district nonrandom "selection" into HOLC A-D grade assignment such that said assignment correlates with unobservable determinants of our outcome variables. While modern-day schools and districts were not assigned HOLC A-D grades and, in many cases, did not even yet exist, selection bias remains a threat if past HOLC A-D neighborhood grade assignments captured pre-existing inequality between neighborhoods variation (e.g., homeownership rates, home values, access to public services, crime) and is a determinant of our modern-day educational outcomes. Notably, Aaronson et al. (2020) find evidence that this is indeed the case while also presenting causal evidence in Aaronson et al. (2021). As indicated in the model specification, we do not use additional covariates in these models. We explored using covariates for family income, percent Black, and building age available for a smaller sample of HOLC polygons as produced by Markley (2021) and did not find the incorporation of these covariates changed our findings. Therefore, we chose the models without covariates to preserve sample size.[12] Still, given concerns about selection bias, we avoid framing our results as causal and instead focus on describing the historical long-term relationship between HOLC A-D grades and

[12] For finance outcomes *s = 6,553* versus 8,573 without covariates; for racial diversity outcomes s = 6,553 versus 9,709 without covariates; for student performance outcomes s = 3,906 versus 5,124 without covariates

their conceptions of neighborhood quality and modern-day educational outcomes.

For the student diversity outcomes, in addition to raw differences in percentages of different racial groups, we also analyze each school's Simpson's Diversity Index (1-D) for our within-school-between-student racial diversity outcome. This index captures the likelihood that two randomly selected students from a given school will belong to different racial groups, ranging from 0 to 1, with larger values representing greater within-school-between-student racial diversity (Simpson, 1949; Hirschman, 1964).

6 Results

We present our results by the three overarching outcomes addressed in this paper: (1) district and school-level finance, (2) school-level racial diversity, and (3) school-level student performance.[13] In each section, we provide and discuss the regression results from our CBSA FE estimation model described in the "Empirical Strategy" section above. For the district finance and racial diversity outcomes, we also provide a time series analysis that looks at how, if at all, the relationship between said outcomes and HOLC A-D security ratings changed over the past three decades.[14]

6.1 District and School Finance

6.1.1 Current Outcomes – Districts

In Table 2 Panel A, we provide nationwide results of our district-finance outcomes. We see that, on average, total per-pupil revenues are often lower for districts mapped to HOLC D grades than those mapped to A, B, and C grades. Nationwide, districts mapped to HOLC D grades have, on average, $14,402 per-pupil total revenues with differences greatest between A

[13] There is notably a mismatch in school years between the district and school spending data and SEDA data. However, the consequences of having test score data on fourth graders in 2009 and spending data for schools and districts a decade later are immaterial. We find comparable results using the SY2009-10 F-33 finance data (adjusted to USD 2018) in the magnitude of estimates and statistical significance.
[14] Our time series analysis does not include SEDA school-level student achievement outcomes or NERDS school-level finance outcomes, as neither contains data from the 1980s or 1990s.

vs. D districts ($1,546, SE = $822) and least between B vs. D districts ($992, SE = $536). All these differences are statistically significant. At the federal level, districts mapped to HOLC D grades have, on average, $1,201 in per-pupil federal revenues, with A (-$722, S.E. = $139) and B (-$391, S.E. = $87) districts being outstripped in federal funding by D districts, and C districts ($253, S.E. = $76) just edging out D districts. Once again, all these differences are statistically significant. Moving to the state revenues, we find that D districts have, on average, $7,960 in per-pupil state funding, with differences largest between A vs. D districts ($3,670, S.E. = $752) and smallest between C vs. D districts (-$461, S.E. = $495). All but the C vs. D average comparisons are statistically significant.

Finally, at the local level, we see that across the U.S., districts mapped to HOLC D grades have, on average, $5,240 in per-pupil local revenues, which is less than those districts mapped to HOLC A, B, and C grades by -$5,937, -$4,592, and -$1,713, respectively. All differences are statistically significant. A cohesive narrative emerges by combining these results with our above state, federal and total results. Districts mapped to HOLC D grades have significantly less per-pupil local funding than higher-rated HOLC A, B, and C districts in the sample. This gap in local funding is abated by redistributive federal and state funds that favor D mapped districts relative to A, B, and C districts; however, non-local educational funding is not enough to overcome the sizeable initial gap driven by differences at the local level.

6.1.2 Longitudinal Outcomes - Districts

We expand our analysis to look at how, if at all, the relationship between HOLC A-D grades and the finance outcomes changed over time. This exercise reduces our original cross-sectional district-level analysis sample (d = 1,760) by a third (d_p = 1,109).[15]

[15] To ensure those districts that remain in our panel dataset are representative of those contained in the original 2017-18 cross-sectional sample, we perform robustness checks across samples. While the distribution of HOLC A-D grades across districts is slightly different between the 2017-

In Figure 2, we show A-D averages across time for each educational finance outcome, weighted by district enrollment. Each outcome is adjusted for inflation and denominated in 2018 USD. Starting with average per-pupil total revenue, one can see near-parallel lines across time with only marginal differences in slopes by HOLC A-D grade. In addition, compound annual growth rates (CAGRs), calculated from 1989-90 to 2017-18, differ little by HOLC security rating and hover around four percent regardless of HOLC grade. While equality in growth rates across time is encouraging, it is less so after considering level differences in average per-pupil total revenue between D districts and their A, B, and C counterparts (Table 2 Panel B). Here we see an initial and subsequently increasing gap in per-pupil total revenue that favors districts located in historically non-redlined neighborhoods (i.e., A, B, C). Thus, while growth rates are similar across HOLC A-D security ratings, per-pupil funding gaps are not. This result is a direct consequence of initial per-pupil total revenue gaps by HOLC A-D grade in the late 1980s paired with near-identical A-D growth rates across time. Findings for per-pupil local revenue mirror those for per-pupil total revenue but are even more pronounced, with larger initial funding gaps in the late 1980s and smaller growth rates.

In contrast to per-pupil total and local revenue trends, per-pupil state and federal revenues favor those districts with lower HOLC security ratings (i.e., C, D) than those with higher (i.e., A, B). Like our 2017-18 cross-sectional findings, state and federal redistributive policies appear to benefit those districts most in need and have consistently done so over the three decades we consider here. For per-pupil state revenues, funding gaps in the late 1980s favor C and D districts relative to A and B districts. Growth rates lead to C and D district convergence, with B districts

---

18 cross-sectional and panel samples, we find no evidence suggesting that average outcomes by A-D HOLC grades vary (Figure A.5, Table A.4). For each F-33 finance outcome, we consider in this paper, we fail to reject the null hypothesis of equality of common coefficients across models using (1) the full cross-sectional 2017-18 sample, and (2) the partial panel 2017-18 sample, where each regression model consists of a given finance outcome regressed on an A-D HOLC indicator variable. This result suggests no statistically significant differences between the weighted average A-D HOLC grades across these two samples for our F-33 educational finance variables.

remaining mostly in parallel; however, there is a distinct increase in the gaps between B, C, and D districts and their highest-rated A counterparts. These results are further confirmed in Table 3, where there is limited variation across years for D vs. C and D vs. B comparisons but a clear monotonic increasing relationship for the D vs. A group. Finally, per-pupil federal revenues match the findings for per-pupil state revenues, with notable differences, including the large positive spike in per-pupil federal funding for SY2009-10 consistent with the surge in education funding from the American Recovery and Reinvestment Act of 2009 (ARRA) (U.S. Department of Education, 2009).

[Insert Table 2 here]

[Insert Figure 2 here]

6.1.3 Current Outcomes – Schools

In contrast to the district finance results, at the school level, we find that, on average, schools mapped to HOLC D schools have more per-pupil total expenditures than schools mapped to A, B, and C HOLC grades (Table 3). Differences between D schools and their higher HOLC A, B, and C schools are monotonically decreasing in HOLC security rating, such that gaps in per-pupil total expenditures are widest between schools mapped to A and D HOLC grades (-$1,539, S.E. = $852) and smallest between schools mapped to C and D HOLC grades (-$882, S.E. = $334). These differences are all statistically significant.

While the 2018-19 NERDS data affords us school-level finance data, it is more coarsely disaggregated relative to the NCES F-33 district finance data, limiting our ability to deconstruct the underlying variation per-pupil total expenditures into cleanly delineated federal, state, and local buckets. That said, we can isolate per-pupil federal expenditures and investigate the combined sum of state and local expenditures. Like the district-level finance results, we find that, on average, schools mapped to D schools have more per-pupil federal expenditures than A, B, and C schools. These differences are all statistically significant and once again are largest between A and D schools (-$602, S.E. = $39) and smallest between A and C schools ($141, S.E. = $31).

[Insert Table 4 here]

For the combined state and local per-pupil expenditure outcome, we observe similar patterns in the average differences between D vs. A, B, C schools as we do for total and federal per-pupil expenditures; however, not all differences are statistically significant, namely, A vs. D differences (-$1,131, S.E. = $860). That said, average differences between D vs. B (-$1,010, SE = $590) and C (-$745, SE = $323) schools are statistically significant. Negative per-pupil state and local averages mean one of two things – per-pupil local expenditures outpace per-pupil state expenditures, or per-pupil state expenditures outpace per-pupil local ones. Incorporating what we know from the F-33 district finance results, where per-pupil state revenues offset large and statistically significant differences in local per-pupil revenues that favor those districts mapped to higher HOLC A, B, or C grades relative to those mapped to redlined HOLC D grades, we argue the latter is most realistic.

6.1.4 Reconciliation of District and School Finance Results

In this section, we attempt to reconcile the differences in the district and school-level results by proposing hypotheses and discussing their likelihood. Starting with the first, the differences in the district and school-level results may simply be an artifact of the outcomes chosen since the district-level analysis uses per-pupil revenues and the school-level analysis uses per-pupil expenditures.[16] If notable variation exists between funding allocations versus reported

[16] We use NCES F-33 revenues because it allows us to explore total revenues and the constituent local, state, and federal revenues that roll up to it. NCES F-33 expenditures do not provide this level of detail.

spending in non-uniform ways by HOLC A-D security rating, one would expect to see these differences play out in our regression results. When we substitute district per-pupil total expenditures for district per-pupil total revenues, our district-level results are indistinguishable from another, indicating that measurement differences are an unlikely driver of our varying school and district finance results.

Another explanation is the notable differences in the school and district mapping strategies we use to assign HOLC grades. Recall, for districts, HOLC A-D mapping is based on the weighted average of the portions of HOLC A-D polygons contained within that district; however, for schools, HOLC A-D mapping is based on a one-to-one match of an individual school to a HOLC grade. By construction, a district with a given HOLC grade assignment will contain the largest share of that HOLC polygon type, in square miles, relative to other HOLC polygons; however, this does not preclude other HOLC graded polygons from being present in that district. If schools are uniformly distributed throughout a district, we would expect within-district school mappings to reflect the within-district distributions of HOLC polygons, but if schools cluster in HOLC polygons different from the district HOLC A-D mapping, the distribution of HOLC school mappings will not accurately reflect the HOLC polygons that make up that district.[17]

[Insert Figure 3 here]

To check this, we construct the underlying HOLC A-D school distributions of each HOLC A-D district. In Figure 2, notable patterns emerge. First, higher (lower) rated HOLC districts have larger shares of higher (lower) rated HOLC schools, with the greatest being that of

[17] For example, suppose a district is composed of mostly HOLC A polygons but has small HOLC B-D polygons where schools cluster. Based on our district and school mapping strategies, the district will be assigned a HOLC A grade, whereas schools will be assigned B-D grades. A scenario such as this could explain our district and school-level finance results.

the HOLC A-D school rating congruent with the HOLC A-D district rating. For example, districts mapped to HOLC A security ratings have the greatest share of HOLC A schools in that district. Alternatively, districts mapped to HOLC D security ratings have the greatest share of HOLC D schools in that district. These findings hold for B and C districts as well. Second, shares of HOLC A-D schools in HOLC A-D districts follow a strict rank order, such that HOLC D districts have the largest share of D mapped schools, second of C, third of B, and the least of A. Conversely, HOLC A districts have the largest share of A mapped schools, second of B, third of C, and the least of D. This pattern holds for B and C districts. These results validate our mapping strategies, and suggest schools are not clustering in polygons discordant from the mapped district HOLC A-D grade.

A final likely explanation is that districts are simply allocating resources to those most in need (e.g., low-income students), which are those schools located today in what were once HOLC D neighborhoods. We check this by regressing the percentage of free and reduced-price lunch students (FRPL) in a school on HOLC grade indicators using the same CBSA FE model as before. We find that nationally, schools mapped to HOLC D grades have, on average, 79.3 percent of students qualifying for free and reduced-price lunch, where percentage point differences are greatest between A vs. D schools (-0.37, S.E. = 0.02) and smallest between A vs. C schools (-0.06, S.E. = 0.01). We observe a similar pattern across all regions. All differences between D vs. A, B, and C schools are statistically significant nationwide and by region. Combined with the district and school finance results, these findings suggest that districts, regardless of their A-D assigned grade, systematically target and allocate more funding to schools located in historically HOLC D neighborhoods because those schools today serve the largest shares of students from low-income households. These results can be seen as an extension of recent research that finds redlined HOLC D neighborhoods are worse off, both in terms of homeownership rates and home values, relative to their higher-rated peers (Aaronson et al., 2020) and provide some of the first evidence of secondary ripple effects that stem from the adverse impacts HOLC maps had on neighborhood quality and development.

Despite the school and district result differences, we argue the district results are still pertinent, as they represent an upper limit on how much a district can allocate to their most-in-need schools (i.e., D schools) while still maintaining adequate funding levels for all other schools they serve. Given that we find that districts' HOLC A-D mapping is predictive of their total per-pupil funding, such that districts mapped to higher-rated HOLC grades have, on average, more per-pupil total funding than those mapped to lower-rated HOLC grades, one might expect lower-rated HOLC schools in higher-rated HOLC districts to receive and spend more money relative to those in lower-rated HOLC districts. We test this theory by looking at kernel density plots (Figure A.4) and regressing per-pupil total funding HOLC indicators for a subset of our analysis sample containing only D schools.[18] While the kernel density analysis does provide evidence to support this theory, the CBSA FE regression results do not, which is likely an artifact of our smaller subsample and less within-CBSA variation in HOLC A-D districts. Zooming out to the state and regional levels, our results from the state F.E. and region F.E. support this narrative and align with the distributional patterns in Figure 3.[19]

This evidence suggests that districts with more resources in our sample (i.e., A, B, C), relative to those with less (i.e., D), distribute more to their most-in-need schools (i.e., D schools).

---

[18] While there is considerable overlap between the F-33 district finance sample and the NERDS 2018-19 school-level sample, up to this point, our school and district finance samples are not perfectly congruous; that is, every district in the NCES F-33 sample is not in the NERDS 2018-19 sample, and vice versa. We link our NERDS school and NCES F33 district-level samples to facilitate a complete district-school comparison. Overall, these datasets share 95% of the same districts; however, each is plagued with missing outcome data, which after removing, leaves us with s = 6,670 schools that roll up to d = 1,025 unique districts. These samples represent just under 60% and 70% of the original F33 district-level analysis and NERDS school-level analysis samples, respectively.

[19] Results are statistically significant when clustering at the city or state level. Results become statistically insignificant ($p = 0.15$) for D vs. B when clustering at the regional level but remain statistically significant for D vs. C differences.

Thus, while D schools have the largest per-pupil total expenditures of all HOLC A-D schools regardless of HOLC A-D district type, our results show that variation in district resources yields variation in D school resources that favor higher-rated HOLC districts relative to lower-graded HOLC districts. Equalization of district per-pupil funding across HOLC A-D district ratings through targeted local, state, and federal programs could help equalize funding at the school level for those located today in historically redlined HOLC D neighborhoods that serve disproportionately low-income students.

6.2 School Racial Diversity

6.2.1 Current Outcomes

In Table 4 Panel A, we see that nationwide schools with HOLC D grades relative to those schools with A, B, and C ratings have larger shares of Black and Non-White students in their schools. Overall, for those schools mapped to HOLC D grades in our sample, 36 percent (S.E. = 0.011) of students are Black, with differences greatest between A vs. D schools (-0.14, S.E. = 0.027) and smallest between C vs. D schools (-0.08, S.E. = 0.016). All these results are statistically significant. For percent non-White, we observe identical but inverse patterns to our findings above for percent White. Our final school diversity outcome variable is the Simpson's Diversity Index. Nationwide, the Simpson's Diversity Index monotonically decreases in HOLC A-D grade, such that A schools have, on average, the most diverse student populations and D schools have the least (0.42, S.E. = 0.009). Differences between D schools and their higher HOLC A (0.07, SE = 0.024), B (0.05, SE = 0.019) and C (0.03, SE = 0.011) are all statistically significant.

These results point to a more nuanced narrative than, for example, D schools have predominantly Black students, whereas A and B schools have primarily White. While it is true that schools with higher-rated HOLC grades have far fewer shares of Black students than D schools, they also have, on average, greater diversity. Pairing these findings with our above district and school-finance results shows that those schools mapped to higher HOLC security ratings have more student diversity, smaller shares of Black students, and serve more affluent households.

What might be driving these results? Wealthy families are less restricted spatially and less likely to be priced out of neighborhoods that function as gateways to high-performing schools. Thus, across the racial spectrum, high-performing districts attract and retain more affluent families that can afford the price of admission, namely a residential property located in the school district. In addition, if diversity mandates are in place, high-performing schools and districts can more easily recruit and retain families from diverse backgrounds relative to their low-performing counterparts. Under this model, enrollment patterns in high-performing districts would be more responsive to changes in diversity goals. In contrast, enrollment patterns in low-performing districts would be less responsive and more reflective of the neighborhood's status quo demographics. Research shows that historical federal housing policies often buttressed patterns of neighborhood segregation by income and race, such that low-income minority families are, by design, spatially concentrated into a select few locales and therefore clustered in neighborhood schools and districts with high rates of student poverty (Katz & Turner, 2009; Turner & Berube, 2009). Given this, we should expect low-performing schools to be less diverse, have higher concentrations of racial minorities, and have more low-income students. These takeaways are consistent with research on the school and neighborhood diversity in urban areas (e.g., Filardo et al., 2008; Candipan, 2019) and recent research on diversity and family income in U.S. K-12 public schools (U.S. Government Accountability Office, 2016). Consequently, our results are consistent with this theory and lend further evidence to the long and harmful associations of historically discriminatory neighborhood policies with educational outcomes.

6.2.2 Longitudinal Outcomes

In this section, we expand our analysis to look at how, if at all, the relationship between HOLC A-D grades and the racial diversity outcomes changed over time. Like the educational finance time series analysis, we first create a panel dataset that spans three decades and includes school-level student demographic data from the late 1980s to the late 2010s. This exercise reduces our original cross-sectional school-level analysis sample (n = 9,709) by half (n_p = 4,677).[20]

First, to better understand how school student demographics and racial diversity changed over time by HOLC A-D grade, in Figure 4 we show A-D averages and differences across time for each racial diversity outcome, weighted by school enrollment. All mean outcomes and differences are captured in Table 3 Panel B. Starting with the percent Black outcome, overall, we see negative downward sloping convex trends from 1988-89 through 2018-19. For those schools located in historically rated A and B neighborhoods, there is a small uptick in the percent Black from 1988-89 to 1998-99, reaching 24.8 and 26.5 percentage points, but this was overwhelmed by negative trends in the two decades that follow, so much so that both groups end up below their original 1988-89 shares. Positive gaps between D vs. A, B, and C schools exhibit an initial downward trend from 1988-89 to 2008-09 and flatten over the final decade. The gap in average

[20] To check if those schools that remain in our panel dataset are representative of those in the original 2018-19 cross-sectional sample, we perform robustness checks across samples. Overall, the distribution of HOLC A-D grades across schools varies little between the 2018-19 cross-sectional and panel samples. However, we find evidence that average outcomes by A-D HOLC grades differ by sample. While often small, these differences are often statistically significant for A-D HOLC grades across each diversity outcome except for Simpson's Diversity Index, which has only statistically significant differences between samples for the B security rating (Figure A.6, Table A.2). With these details in mind, we discuss the time series results below. For each diversity outcome we consider in this paper, we reject the null hypothesis of equality of common coefficients across models using 1) the full cross-sectional 2018-19 sample, and 2) the partial panel 2018-19 sample, where each regression model consists of a given diversity outcome regressed on an A-D HOLC indicator variable. This result suggests statistically significant differences between the weighted average A-D HOLC grades across these two samples for our diversity outcomes. The exception is the Simpson's Diversity outcome, which fails to reject the null hypothesis of equality of common coefficients for A, C, and D HOLC grades across samples.

shares of Black students is largest between the D vs. A group beginning at 13.4 percentage points in 1988-89 and shrinking to 10.3 percentage points by 2018-19. In comparison, the gap between D vs. B and D vs. C schools begins at 10.8 and 7.8 before decreasing over the following decades to end at 7.4 and 6.5 percentage points, respectively.

We see parallel downward sloping lines across time for the percent White outcome, with negative gaps between D vs. A, B, and C counterparts remained mostly flat from 1988-89 through 2018-19. Thus, although student bodies have become less White over time across all HOLC A-D grades in our time series sample, they have done so at similar rates. There is a clear rank order by HOLC A-D grade to the lines in Figure 4, with A schools having, on average, the largest share of White students, B the second, C the third, and D the fourth and smallest. Thus, gaps are largest between A and D schools, with differences of around 35 percentage points. For the percent Non-White outcome, we see the same patterns as the percent White outcome, except in reverse. Finally, for the Simpson's Diversity (1-D) outcome, we see uniform increases across all HOLC A-D grades from 2008-09 to 2018-19, with imperceptible gaps between A and B schools but notable negative differences between D vs. A, B, and C counterparts. However, these patterns are not consistent across time. For example, in 1988-89, average diversity levels hover around 0.35 across all HOLC A-D grades, with gaps near zero between D vs. A, B, and C schools. From this point forward, the diversity index steadily increases for those schools in the highest-rated HOLC neighborhoods (i.e., A, B) while remaining flat for those schools in the lowest-rated HOLC neighborhoods (i.e., C, D). Thus, by 2008-09 there were notable negative gaps in the diversity index between D vs. A, B, and C schools that continue to grow into 2018-19. This pattern is especially true for D vs. A, B comparisons, where previously near-zero gaps in the diversity index in 1988-89 surpass -0.08 in 2018-19.

[Insert Table 3 here]

[Insert Figure 4 here]

This longitudinal analysis shows student racial diversity increasing over time for all HOLC A-D grades. Notably, we see little movement in Simpson's Diversity Index from the late 1980s through the late 2000s for those in historically HOLC D neighborhoods. This might once again reflect the waning influence of court-order desegregation plans starting in the 1990s, and standalone could be a harbinger of resegregation in the years to follow. However, this trend reverses and spikes upward in the last decade, joining already upward sloping trend lines for A, B, and C schools. While these patterns could reflect more recent efforts that target school racial diversity through new avenues such as SES integration instead of historical policies based on race integration (Wells et al., 2020), it may also be an artifact of our chosen diversity measure. Finally, over the three decades we consider, we observe gaps between D vs. A, B, and C grades that are persistent and often grow over time. These inequalities between those schools located today in what were historically the best-rated neighborhoods and those located today in what were the worst-rated neighborhoods highlight the potentially stubborn historical legacy of HOLC A-D map grades.

2.6.3 School Student Performance

2.6.3.1 Current Outcomes

Lastly, we present student performance outcomes. Nationwide and by region, there are virtually no statistically significant differences across HOLC A-D grades for the outcomes of average student learning rates and average student test score trends. When differences are statistically significant, they are small as a share of the total variation in student test scores, with the largest statistically significant difference equating to just over 1/20th of a grade level (Table

4). Each of their respective HOLC A-D grade averages, both nationwide and by region, is not statistically different from their respective grand means. That said, given that the between-school standard deviation average student learning rates (i.e., ~0.07) and average student test score trends (i.e., ~0.04) are also small, the point estimates for these measures constitute a large share of the between-school standard deviation. Looking at these two outcomes, we see this is particularly the case for average student test score trends where differences between HOLC D vs. A (0.010, S.E. = 0.003), B (0.007, S.E. = 0.002), and C (0.004, S.E. = 0.001) make up 25%, 18% and 10% of the between-school standard deviation, respectively.

[Insert Table 4 here]

In contrast, the average student Math and ELA score outcome exhibits statistically significant differences between HOLC A-D grades nationwide and across all regions. This is true for all HOLC D schools versus their higher A, B, and C counterparts. Also, moving from A to B, B to C, and C to D, average student test scores decrease monotonically such that the gap between A and D is the widest among all D vs. A, B, and C differences. This is once again true both nationwide and across each region.

These results tell us that while learning rates (i.e., how students' scores improve each school year) and changes in educational opportunity (i.e., trends in test scores within grades across cohorts) are, on average, the same across all HOLC A-D grades, overall educational opportunity (i.e., average students' test scores) is not. Specifically, those schools located in historically D-assigned neighborhoods have less educational opportunity than those in A, B, and C neighborhoods. For example, in Table 4, nationwide A and D schools are separated by 0.64 SD units (S.E. = 0.04) or just over 1.9-grade levels. These gaps are present across all regions and widen to as much as 0.75 SD units (S.E. = 0.05) or about 2.25-grade levels in the West region and shrink to 0.54 SD units (S.E. = 0.10) or around 1.6-grade levels in the Northeast (Table S0.4). Comparing B and D schools paints a similar picture to above, albeit muted, with schools separated by 0.34 SD units (S.E. = 0.03) or just above one grade level. These gaps favoring B versus D schools are also exhibited within each region, increasing to as much as 0.46 SD units (S.E. = 0.07) in the West and decreasing to as little as 0.30 SD units (S.E. = 0.06) in the South (Table S0.4). Finally, C and D schools show the greatest similarity of all D vs. A, B, and C comparisons, with gaps shrinking to single digits nationwide and across regions. Overall, C and D schools are separated by 0.13 SD units (S.E. = 0.02) or just under one-half grade level. Differences in educational opportunity are largest for the South region with a gap of 0.15 SD units (S.E. = 0.05), while the Northeast gap is smallest at 0.10 SD units (S.E. = 0.02) (Table S0.4). Both nationwide and by region, these differences in D vs. A, B, and C educational opportunity are often statistically significant.

What do these findings mean? First and most importantly, they are only a snapshot of present-day differences in educational performance measures by HOLC A-D grade. They thus cannot speak to whether gaps have risen or fallen over the past several decades. Without a complete accounting of these trends over time, we cannot measure progress nor bring historical context to bear. For example, suppose trends in educational opportunity favored D vs. A, B, and C schools over the past half-century. In that case, one might view the current gaps in this outcome as a historical lower bound and vice versa for a historical upper bound. Unfortunately, apart from the current 2009-2018 SEDA panel, we lack historical data on educational performance measures. Even so, we can make prognostications on what might be if the status quo remains. Given the large gaps in educational opportunity by HOLC A-D security rating and the near-zero S.D. unit changes in it for each HOLC A-D grade, the educational opportunity gap is expected to remain unabated into the future. The equality exhibited in average learning rates and average educational opportunity changes by HOLC A-D grade, which standalone might be a positive finding, will lead to a continued inequality in average educational opportunity across them, given the large and existing gaps in educational opportunity by HOLC A-D grade.

Finally, this equilibrium could have a positive aspect if school and later life outcomes only depend on meeting a minimum educational opportunity threshold. For example, while gaps in educational opportunity would remain constant over time, a positive average change in educational opportunity uniform across A-D HOLC grades could, in time, raise all schools to and above the minimum educational opportunity threshold. Unfortunately, our findings do little to support this claim, as changes in educational opportunity, while positive, are small and often fall below 0.01 SD units.

2.7 Discussion and Conclusion

Between 1935-1940, the Home Owners' Loan Corporation (HOLC) assigned A (minimal risk) to D (hazardous) grades that arguably had meaningful effects on how the FHA, private banks, and mortgage lenders evaluated the creditworthiness and risk of home loans and mortgage insurance within residential neighborhoods over the next several decades. With the release of newly digitized HOLC A-D maps from the University of Richmond lead "Mapping Inequality Project," there has been a recent surge in research quantifying the negative impacts of redlining on long-term social and economic outcomes (Appel & Nickerson, 2016; Krimmel, 2018; Anders, 2018; Aaronson et al., 2020; Aaronson et al., 2021). However, this effort has yet to extend to K-12 public school educational outcomes to the best of our knowledge.

This paper examines the relationship between historic HOLC A-D maps and modern-day district and school funding patterns, racial diversity, and student performance. We employ a novel mapping strategy that links 1935-1940 HOLC A-D neighborhood grades to present-day districts and schools. At the district level, we find those mapped to historic HOLC D grades have the least favorable overall and local district-finance outcomes relative to those mapped to higher-rated HOLC A, B, and C grades.[21] These results show how inequality in local per-pupil funding drives inequality in total per-pupil funding at the district level. Our findings also highlight the mitigating effects of redistributive federal and state policies on funding gaps generated by local differences. For example, we find those districts mapped to historic HOLC D grades have the most favorable state and federal district-finance outcomes today. These findings show a redistributive system targeting districts with higher percentages of students eligible for free or reduced-price lunch. However, the results also suggest that past neighborhood inequality lingers well into the future. For example, those districts that receive redistributive funding to equalize local funding inequities are also those that serve families in neighborhoods disproportionately composed of HOLC D grade polygons.

In contrast to our district finance findings, we find the inverse relationship between HOLC A-D grades and funding at the school level. Schools mapped to the worst HOLC grades (i.e., D) have the most favorable school-finance outcomes than their higher-rated counterparts (i.e., A, B, and C). In reconciling the district and school-level finance results, we find that D schools have the greatest share of students eligible for free or reduced-price lunch relative to A, B, and C schools. Together, the district and school-level findings suggest that districts, regardless of their HOLC A-D grade, systematically target and allocate more money to schools in historic HOLC D neighborhoods and do so because these schools serve the largest shares of students from low-income households. Finally, these results persist across time, with overall positive time

[21] Recall, that discrete A-D HOLC grades for districts are based on A-D HOLC weighted averages where weights were derived from A-D HOLC grade polygon areas.

trends in outcome measures regardless of HOLC A-D grade but widening gaps between D vs. A, B, and C districts.

Overall, these results align with recent research that finds HOLC negatively impacted the development of urban neighborhoods and led to lower homeownership rates, home values, and racial diversity decades later (Aaronson et al., 2020; Aaronson et al., 2021). Making out-of-sample predictions from Aaronson et al. (2021) suggest that districts composed primarily of HOLC D neighborhoods (and therefore mapped to HOLC D grades in our paper) should have lower assessed property values, lower property tax bases, and less local funding than those districts composed primarily of A, B, and C HOLC neighborhoods. In general, any local funding shortfalls are addressed through targeted state and federal redistributive funding programs (e.g., Title I, IDEA) that allocate dollars to low-income districts to reduce funding gaps and better equalize financing across districts. Our district finance results support this narrative and highlight a state and federal funding apparatus that is effectively targeting those districts most in need, albeit at insufficient amounts to equalize funding altogether. While our results are not causal, our hypothesized mechanism underlying them hints at a lingering historical legacy of redlining, where HOLC neighborhood grades assigned in 1935-1940 predict local funding gaps today and where federal and state funds are needed to equalize funding for districts composed of primarily D neighborhoods.

We also find those schools located today in historically redlined residential neighborhoods to have, on average, larger shares of Black and non-White student bodies and less diverse student populations. These differences are persistent and growing over time for the racial diversity outcomes, albeit for a smaller, less representative sample. However, while A assigned public schools have the highest percent White student populations, they also exhibit the highest student racial diversity levels via Simpson's Diversity Index. That said, A schools have the lowest Exposure Index values across all HOLC A-D grades and race-group pairings (i.e., White-Asian, White-Black, White-Hispanic, White-Non-White).

These findings reflect broader trends in U.S. K-12 public school demographics that have led to today's more racially and ethnically diverse school-age population. For example, over the past two decades, shares of White 5-17-year-olds have decreased from 62 percent to just over 50 percent, whereas shares of Hispanic 5-17-year-olds have increased to 25 percent from 16 percent (NCES, 2019). These public-school demographic shifts mirror the increasing racial diversity of the U.S. population, driven in part by diversifying urban demographics resulting from nationwide migration patterns that brought White families into cities from the suburbs and Black, Hispanic, and Asian families out to them (Wells et al., 2020). Even so, more diverse populations may not always translate to more diverse schools. Since the 1990s, court-ordered desegregation plans from the 1960s and 1970s have been gradually lifted, leading to increased school segregation (Lutz, 2011; Reardon et al., 2012; Reardon et al., 2019). Also, intergroup exposure between Whites and Non-White students has decreased since the 1990s, with Non-White students attending schools with fewer shares of White students (Fiel, 2013).[22] These countervailing forces could overwhelm, or at a minimum, limit the benefits that a more diverse U.S. population has on school racial diversity.

Finally, we also find that schools located today in what once were historically redlined areas have worse average ELA and math scores. However, there is no difference in both average learning rates and trends in test scores across A-D schools. Notably, these findings only provide

[22] Importantly, Fiel (2013) finds this result was due to a growing share of the minority population relative to whites, not from increasing between-group segregation. This is reflected in the negative trends in percent White and positive trends in percent Hispanic and percent Non-White we present in this paper.

a snapshot of present-day differences in educational performance measures by HOLC A-D grade. They do not lend insight into whether gaps by HOLC A-D grade have abated over time. That said, looking forward, if the status quo continues, we will predict the gap in average test scores by HOLC A-D grade to remain unabated, given the limited to no difference in average learning rates and average test score trends by HOLC A-D grade.

Overall, our paper provides evidence that shows the stubborn association of HOLC A-D maps with modern educational outcomes and highlights the transmission of past neighborhood inequality to the present. In addition, these results suggest that education policymakers need to consider the historical implications of past neighborhood inequality on present-day neighborhoods when designing and implementing complex modern interventions that target inequitable outcomes between students of different socioeconomic and racial groups.

References

Aaronson, D. (1998). Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes. *The Journal of Human Resources,* 33(4), 915–946.

Aaronson, D., Hartley, D., Mazumder, B. (2020). The Effects of the 1930s HOLC "Redlining" Maps. *Federal Reserve Bank of Chicago*. Working Paper.

Aaronson, D., Faber, J., Hartley, D., Mazumder, B., & Sharkey, P. (2021). The long-run effects of the 1930s HOLC "redlining" maps on place-based measures of economic opportunity and socioeconomic success. *Regional Science and Urban Economics*, 86, 103622, 1-15.

Alcaino, M., & Jennings, J. (2020). How Increased School Choice Affects Public School Enrollment and School Segregation. *EdWorkingPaper* No. 20-258. Providence, RI: Annenberg Institute at Brown University. Retrieved December 8, 2020, from https://www.edworkingpapers.com/ai20-258.

Anders, J. (2018). The Long Run Effects of De Jure Discrimination in the Credit Market: How Redlining Increased Crime. Working Paper. Retrieved March 1, 2021, from https://johnanders625665825.files.wordpress.com/2018/12/anders_redlining_12_16_2018.pdf

Appel, I., Nickerson, J. (2016). Pockets of Poverty: The Long-Term Effects of Redlining. *SSRN Electronic Journal*, SSRN Electronic Journal.

Baker, B.D., Farrie, D., Sciarra, D., Luhm, T. (2018) Is School Funding Fair? 2018 Edition. Education Law Center of New Jersey & Rutgers GSE.

Bayer, P., Blair, P., & Whaley, K. (2020). A National Study of School Spending and House Prices. NBER Working Paper No. w28255. Cambridge, MA: National Bureau of Economic Research. Retrieved March 1, 2021, from https://www.nber.org/papers/w28255.

Beard, J. H., Morrison, C. N., Jacoby, S. F., Dong, B., Smith, R., Sims, C. A., & Wiebe, D. J. (2017). Quantifying disparities in urban firearm violence by race and place in Philadelphia, Pennsylvania: A cartographic study. *American Journal of Public Health*, 107(3), 371-373.

Brittain, J., Willis, L., & Cookson, P. (2019). Sharing the Wealth: How Regional Finance and Desegregation Plans Can Enhance Educational Equity. *EdWorkingPaper* No. 19-187. Providence, RI: Annenberg Institute at Brown University. Retrieved March 1, 2021, from https://www.edworkingpapers.com/index.php/ai19-187.

Candipan, J. (2019). Neighbourhood change and the neighbourhood-school gap. *Urban Studies (Edinburgh, Scotland)*, 56(15), 3308–3333.

Chetty, R, Friedman, J., Hendren, N., Jones, M., & Porter, S. (2018). The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility. *NBER Working Paper* No. w25147. Cambridge, MA: National Bureau of Economic Research. Retrieved January 5, 2021, from https://www.nber.org/papers/w25147.

Chetty, R., & Hendren, N. (2018a). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *The Quarterly Journal of Economics*, 133(3), 1107-1162.

Chetty, R., & Hendren, N. (2018b). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. *The Quarterly Journal of Economics*, 133(3), 1163-1228.

Chetty, R., Hendren, N., & Katz, L. (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4), 855-902.

Clotfelter, C.T., Hemelt, S.W., Ladd, H.F., & Turaeva, M. (2019). School segregation in the era of immigration, school choice and color-blind jurisprudence – the case of North Carolina. *EdWorkingPaper* No. 19-101. Providence, RI: Annenberg Institute at Brown University. Retrieved March 1, 2021, from https://www.edworkingpapers.com/ai19-101.

Clotfelter, C.T., Ladd, H.F., Clifton, C.R., & Turaeva, M.R. (2021). School Segregation at the Classroom Level in a Southern 'New Destination' State. *Race and Social Problems,* 13(2), 131-160.

Connolly, N. (2014). A World More Concrete: Real Estate and the Remaking of Jim Crow South Florida. University of Chicago Press.

Crossney, K. & Bartelt, D. (2005). The legacy of the home owners' loan corporation. *Housing Policy Debate*, 16(3-4), 547–574.

Dalane, K., & Marcotte, D.E. (2020). The Segregation of Students by Income in Public Schools. *EdWorkingPaper* No. 20-338. Providence, RI: Annenberg Institute at Brown University. Retrieved March 1, 2021, from https://www.edworkingpapers.com/ai20-338.

Edunomics Lab (2021). NERD$: National Education Resource Database on Schools (Version 1.0). *Georgetown University*. Retrieved 06/2021 from https://edunomicslab.org/nerds/

Fiel, J.E. (2013). Decomposing School Resegregation: Social Closure, Racial Imbalance, and Racial Isolation. *American Sociological Review*, 78(5), 828-848.

Filardo, M., Allen, M., Huvendick, N., Sung, P., Garrison, D., Turner, M. A., Comey, J., Williams, B., & Guernsey, E. (2008). Quality Schools, Healthy Neighborhoods, and the Future of DC: Policy Report. In *21st Century School Fund*. 21st Century School Fund.

Fishback, P., LaVoice, J., Shertzer, A., Walsh, R. (2020). Race, Risk, and the Emergence of Federal Redlining. *NBER Working Paper* No. w28146. Cambridge, MA: National Bureau of Economic Research. Retrieved March 1, 2021, from http://www.nber.org/papers/w28146.pdf.

Green, P., Baker, B., & Oluwole, J. (2021). School Finance, Race, and Reparations. *Washington and Lee Journal of Civil Rights and Social Justice* (forthcoming). Retrieved January, 15, 2021, from https://ssrn.com/abstract=3766279.

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

Hirschman, A. O. (1964) The paternity of an index. *The American Economic Review*, 54(5), 761-762.

Hoffman, J. S., Shandas, V., & Pendleton, N. (2020). The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas. *Climate*, 8(1), 12.

Jackson, K. T. (1980). Race, Ethnicity, and Real Estate Appraisal. *Journal of Urban History,* 6(4), 419–452.

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

Katz, B. & Turner, M.A. (2009). Rethinking U.S. Rental Housing Policy. In *Opportunity 08* (2nd ed., p. 198). Brookings Institution Press.

Krimmel, J. (2018). Persistence of Prejudice: Estimating the Long Term Effects of Redlining. *SocArXiv*, Center for Open Science.

Light, J. (2010). Nationality and Neighborhood Risk at the Origins of FHA Underwriting. *Journal of Urban History*, 36(5), 634-671.

Lutz, B. (2011). The end of court-ordered desegregation. *American Economic Journal. Economic Policy*, 3(2), 130-168.
Markley, S. (2021, September 25). Tabulating HOLC Area Description Sheet Data. https://doi.org/10.31235/osf.io/dktah
Massey, D., & Denton, N. (1988). The Dimensions of Residential Segregation. *Social Forces*, 67(2), 281-315.
Massey, D., & Denton, N. (1993). American apartheid: Segregation and the making of the underclass (Democracy and urban landscapes). Cambridge, MA: Harvard University Press.
Mitchell, B., & Franco, J. (2018). HOLC Redlining Maps: The Persistent Structure of Segregation and Economic Inequality. Washington, DC: National Community Reinvestment Coalition. Retrieved March 1, 2021, from https://ncrc.org/wp-content/uploads/dlm_uploads/2018/02/NCRC-Research-HOLC10.pdf.
Monarrez, T., Kisida, B., & Chingos, M. (2020). The Effect of Charter Schools on School Segregation. *EdWorkingPaper* No. 20-308. Providence RI: Annenberg Institute at Brown University. Retrieved March 1, 2021, from https://www.edworkingpapers.com/ai20-308.
Murnane, R.J. (2021). Can policy interventions reduce inequality? Looking beyond test scores for evidence. William T. Grant Foundation.
Nardone, A. L., Casey, J. A., Rudolph, K. E., Karasek, D., Mujahid, M., & Morello-Frosch, R. (2020). Associations between historical redlining and birth outcomes from 2006 through 2015 in California. PloS One, 15(8), E0237241.
Nelson, R.K., Winling, L., Marciano, R., Connolly, N., et al. (2021). Mapping Inequality. American Panorama. Retrieved October 2, 2020, from https://dsl.richmond.edu/panorama/redlining.
NYC Department of City Planning. (2020). School Districts (Clipped to Shoreline). Retrieved December 4, 2020, from https://www1.nyc.gov/site/planning/data-maps/open-data/districts-download-metadata.page.
Reardon, S.F., Grewal, E.T., Kalogrides, D. & Greenberg, E. (2012). Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools. Journal of Policy Analysis and Management, 31(4), 876-904.
Reardon, S. F., Ho, A. D., Shear, B. R., Fahle, E. M., Kalogrides, D., Jang, H., & Chavez, B. (2021). Stanford Education Data Archive (Version 4.0). Retrieved February 8, 2021, from http://purl.stanford.edu/db586ns4974.
Reardon, S.F., Weathers, E.S., Fahle, E.M., Jang, H., & Kalogrides, D. (2019). Is Separate Still Unequal? New Evidence on School Segregation and Racial Academic Achievement Gaps CEPA Working Paper No.19-06. Stanford, CA: Stanford Center for Education Policy Analysis. Retrieved March 1, 2021, from http://cepa.stanford.edu/wp19-06.
Roza, M., Hagan, K., & Anderson, L. (2020). Variation is the Norm: A Landscape Analysis of Weighted Student Funding Implementation. Public Budgeting and Finance. Retrieved March 1, 2021, from https://doi.org/10.1111/pbaf.12276.
Simpson, E. (1949). Measurement of diversity. Nature, 163, 688.
Turner, M.A., & Berube, A. (2009). Vibrant Neighborhoods, Successful Schools: What the Federal Government Can Do to Foster Both. In *Urban Institute*. Urban Institute.
Urban Institute (2017). How do school funding formulas work? Washington, DC: Urban Institute. Retrieved March 1, 2021, from https://apps.urban.org/features/funding-formulas/.

U.S. Census Bureau. (2020). 2019 TIGER/Line Shapefiles (machine readable data files). Retrieved November 8, 2020, from https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html.

U.S. Census Bureau. (2021). Public School Spending Per Pupil Increases by Largest Amount in 11 Years. Retrieved June 8, 2021, from https://www.census.gov/newsroom/press-releases/2021/public-school-spending-per-pupil.html.

U.S. Department of Education. (2009). The American Recovery and Reinvestment Act of 2009: Saving and Creating Jobs and Reforming Education. Retrieved March 1, 2021, from https://www2.ed.gov/policy/gen/leg/recovery/implementation.html

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2019). Status and Trends in the Education of Racial and Ethnic Groups 2018. Publication no. NCES 2019-038. Washington, DC: National Center for Education Statistics. Retrieved February 25, 2020, from https://nces.ed.gov/pubs2019/2019038.pdf.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2020). Retrieved October 2, 2020, from https://nces.ed.gov/programs/edge/Home.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2021). Retrieved December 11, 2021, from https://nces.ed.gov/ccd/elsi.

U.S. Government Accountability Office. (2016). K-12 Education: Better Use of Information Could Help Agencies Identify Disparities and Address Racial Discrimination. Retrieved October 2, 2020, from https://www.gao.gov/products/gao-16-345.

Wells, A.S., Fox, L., Cordova-Cobo, D., & Kahlenberg, R.D. (2016). How racially diverse schools and classrooms can benefit all students. The Education Digest, 82(1), 17.

Woods, L. L. (2012). The Federal Home Loan Bank Board, Redlining, and the National Proliferation of Racial Lending Discrimination, 1921–1950. *Journal of Urban History*, 38(6), 1036–1059.

Acknowledgments: We thank Joshua Goodman, Eric Taylor, Martin West, Andrew Ho, David Deming, HGSE Measurement Lab, and Bridges Collaborative for their feedback on prior drafts and presentations of this project. We acknowledge the financial support from the Bill & Melinda Gates Foundation and the Institute of Education Sciences.

CrediT Authorship Contribution Statement

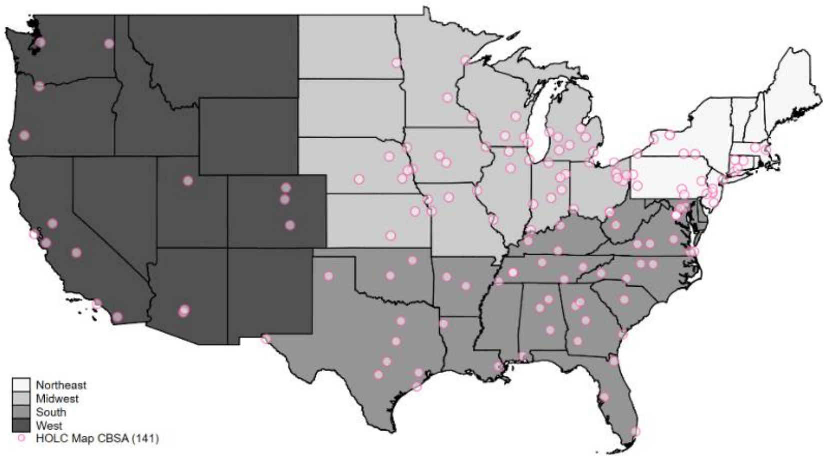
Christopher H. Cleveland: Coequal first author; Conceptualization, Writing – original draft, Writing – review & editing, Project Administration – funding identification & grant application.

Dylan J. Lukes: Coequal first author; Data Curation, Formal Analysis, Methodology, Project Administration – funding acquisition & grant application, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Funding: This work was supported by the Bill & Melinda Gates Foundation, Seattle, WA [INV-036635]. This work was also supported by the Institute of Education Sciences, Washington, DC [R305B150010].

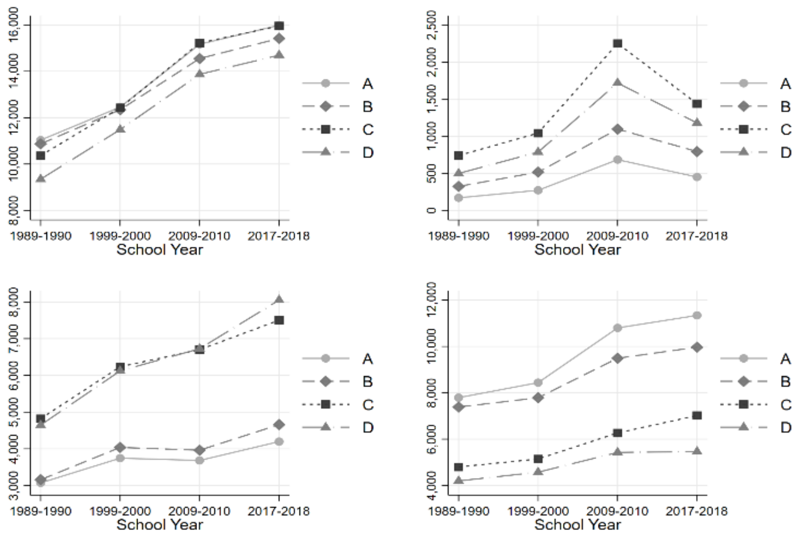
Tables & Figures

Figure 1: 1935-1940 HOLC CBSAs by US Census Bureau Region



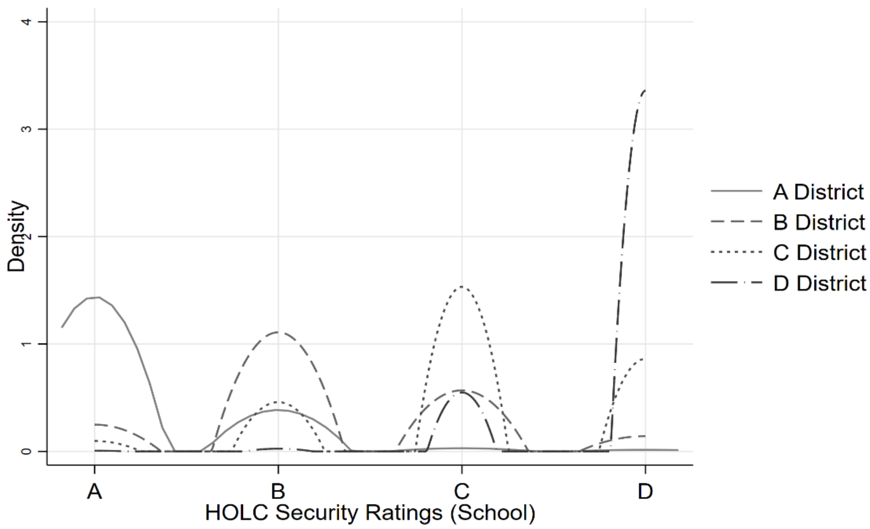
Notes: The above pink dots represent each unique CBSA in our analysis sample. These total to n =144 present-day CBSAs mapped to the 1935-1940 HOLC Residential Security Maps and are broken down by regions as follows: Northeast (n=30), Midwest (n = 54), South (n = 45), and West (n = 15).

Figure 2: HOLC Averages, Finance Outcomes, 1989-2018 (USD 2018)



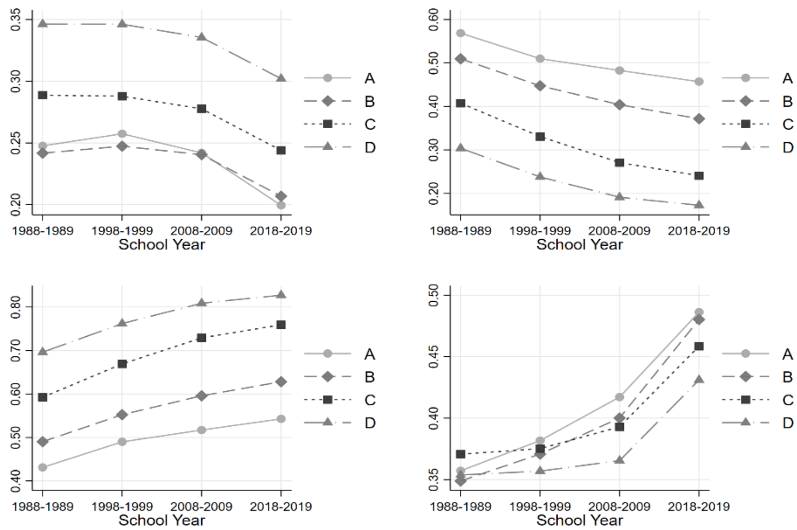
Notes: [top left] Per-Pupil Total Revenue; [top right] Per-Pupil Federal Revenue; [bottom left] Per-Pupil State Revenue; [bottom right] Per-Pupil Local Revenue. All values represented in USD 2018. Weighted averages are from a regression of a given finance outcome on HOLC A-C indicators with student enrollment as an analytic weight. Standard errors are heteroskedasticity-consistent and clustered at the city-level. All dollars denominated in USD 2018.

Figure 3: HOLC A-D School Distributions by HOLC A-D District Grades



Notes: This figure shows the density of HOLC A-D school distributions within the average District grade in which the schools are organized.

Figure 4: HOLC A-D Averages Over Time, Racial Diversity Outcomes, 1988-2019



Notes: [top left] Percent Black; [top right] Percent White; [bottom left] Percent Non-White; [bottom right] Simpson's Diversity Index (1-D). Weighted averages are from a regression of a given diversity outcome on HOLC A-C indicators with student enrollment used as analytic weights. Standard errors are heteroskedasticity-consistent and clustered at the city-level.

Table 1: Summary of Analysis Samples by Outcome Groups

	Fiscal		Student Racial Diversity		Student Performance
	Cross-Sectional	Time Series	Cross-Sectional	Time Series	Cross-Sectional
Outcomes	Per-Pupil Revenue – Total, Local, State, Federal		% Black, White, Non-White, and Simpson's Diversity Index		SEDA Outcomes
Level	District & School		School		School
Year	2017-2018	1989-2018	2018-2019	1989-2018	2009-2018
Schools	N/A		9,709	4,677	5,124
Districts	1,760	1,109	1,955	590	1,006
CBSAs	144	143	141	118	137
District Matches					
F33	1,760	1,109	1,283	581	891
Racial Diversity	1,283	581	1,955	590	1,001
SEDA	891	N/A	1,001	N/A	1,006

Notes: For a more detailed breakdown of outcome group samples please refer to Appendix Tables A.1-A.3. These tables cover shares of schools, districts, and CBSAs included in the study and a variety of characteristics (e.g., urban locale, demographics) of the units represented in our study. The time series sample is not applicable for the student performance outcome group. The count of schools is not applicable for "Fiscal" outcomes since unit of analysis is at the district-level.

Table 2: District Per-Pupil Revenues and 1935-1940 HOLC A-D, SY2017-18 (USD 2018)

Panel A: 2017-2018

Nationwide	District PPR Total	District PPR Federal	District PPR State	District PPR Local
A	$1,546*	-$722***	-$3,670***	$5,937***
	[$822]	[$139]	[$752]	[$1,228]
B	$992*	-$391***	-$3,210***	$4,592***
	[$536]	[$87]	[$776]	[$892]
C	$1,504***	$253***	-$461	$1,713***
	[$514]	[$76]	[$495]	[$468]
Constant	$14,402***	$1,201***	$7,960***	$5,240***
	[$396]	[$62]	[$438]	[$425]
FE	Y	Y	Y	Y
N	1,760	1,760	1,760	1,760

Panel B: Margins & Differences Over Time, Finance Outcomes, 1989-2018 (USD 2018)

		Margins				D vs. A, B, and C		
		A	B	C	D	A vs. D	B vs. D	C vs. D
District PPR – Total (USD 2018)	1989-90	$11,035	$10,870	$10,364	$9,347	$1,688***	$1,524***	$1,017***
		[$439]	[$445]	[$95]	[$265]	[$495]	[$427]	[$331]
	1999-00	$12,457	$12,343	$12,426	$11,486	$971	$857**	$940***
		[$611]	[$380]	[$83]	[$259]	[$668]	[$401]	[$311]
	2009-10	$15,168	$14,552	$15,221	$13,871	$1,297*	$681	$1,349***
		[$693]	[$579]	[$136]	[$298]	[$718]	[$493]	[$404]
	2017-18	$15,989	$15,419	$15,961	$14,698	$1,291	$721	$1,263**
		[$748]	[$630]	[$156]	[$411]	[$873]	[$539]	[$532]
District PPR – Federal (USD 2018)	1989-90	$170	$324	$740	$499	-$329***	-$175***	$241***
		[$58]	[$43]	[$11]	[$50]	[$70]	[$55]	[$59]
	1999-00	$272	$516	$1,042	$786	-$515***	-$270***	$256***
		[$85]	[$60]	[$16]	[$61]	[$93]	[$65]	[$73]
	2009-10	$684	$1,096	$2,251	$1,717	-$1,033***	-$622***	$534***
		[$169]	[$143]	[$32]	[$104]	[$194]	[$182]	[$119]
	2017-18	$450	$793	$1,437	$1,179	-$729***	-$386***	$258***
		[$143]	[$85]	[$22]	[$74]	[$154]	[$103]	[$88]
District PPR – State (USD 2018)	1989-90	$3,072	$3,159	$4,822	$4,648	-$1,575***	-$1,489***	$174
		[$289]	[$181]	[$46]	[$233]	[$359]	[$268]	[$267]
	1999-00	$3,746	$4,039	$6,237	$6,123	-$2,377***	-$2,084***	$114
		[$413]	[$325]	[$66]	[$216]	[$456]	[$419]	[$239]
	2009-10	$3,684	$3,964	$6,701	$6,727	-$3,043***	-$2,763***	-$26
		[$574]	[$437]	[$103]	[$308]	[$601]	[$531]	[$361]
	2017-18	$4,197	$4,659	$7,505	$8,051	-$3,854***	-$3,392***	-$546
		[$686]	[$567]	[$141]	[$379]	[$735]	[$629]	[$461]
District PPR – Local (USD 2018)	1989-90	$7,793	$7,388	$4,802	$4,201	$3,592***	$3,187***	$601
		[$621]	[$556]	[$116]	[$316]	[$675]	[$561]	[$390]
	1999-00	$8,439	$7,787	$5,147	$4,576	$3,863***	$3,211***	$571
		[$890]	[$505]	[$116]	[$392]	[$939]	[$556]	[$469]
	2009-10	$10,800	$9,492	$6,269	$5,427	$5,374***	$4,066***	$842
		[$1,131]	[$714]	[$166]	[$431]	[$1,215]	[$775]	[$523]
	2017-18	$11,341	$9,967	$7,019	$5,468	$5,874***	$4,499***	$1,551**
		[$1,146]	[$713]	[$165]	[$520]	[$1,312]	[$887]	[$595]
FE		Y	Y	Y	Y	Y	Y	Y
N		1,109	1,109	1,109	1,109	1,109	1,109	1,109

Notes: Cluster-robust standard errors are in parentheses with clustering done at the city-level. For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. 2017-18 total students weight all regressions. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. All dollars denominated in USD 2018. *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 3: School Per-Pupil Expenditures and 1935-1940 HOLC A-D, SY2018-19 (USD 2018)

Nationwide	School PPE Total	School PPE Federal	School PPE State & Local
A	-$1,539* [$852]	-$602*** [$39]	-$1,131 [$860]
B	-$1,313** [$581]	-$351*** [$29]	-$1,010* [$590]
C	-$882*** [$334]	-$141*** [$31]	-$745** [$323]
Constant	$16,890*** [$304]	$1,379*** [$18]	$15,658*** [$297]
FE	Y	Y	Y
N	8,573	8,573	8,573

Notes: Cluster-robust standard errors are in parentheses with clustering done at the city-level. For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. 2017-18 total students weight all regressions. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. All dollars denominated in USD 2018. *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 4: School Racial Diversity and 1935-1940 HOLC A-D Grades, 2018-2019

Panel A: 2018-2019

Nationwide	% Black	% White	% Non-White	Simpson's Diversity (1-D)
A	-0.14***	0.3***	-0.3***	0.07***
	[0.027]	[0.022]	[0.022]	[0.024]
B	-0.11***	0.18***	-0.18***	0.05***
	[0.017]	[0.013]	[0.013]	[0.019]
C	-0.08***	0.06***	-0.06***	0.03***
	[0.016]	[0.01]	[0.01]	[0.011]
Constant	0.36***	0.14***	0.86***	0.42***
	[0.011]	[0.007]	[0.007]	[0.009]
FE	Y	Y	Y	Y
N	9,709	9,709	9,709	9,709

Panel B: Margins & Differences Over Time, Racial Diversity Outcomes, 1988-2019

		Margins				D vs. A, B, and C HOLC		
		A	B	C	*D*	A vs. D	B vs. D	C vs. D
% Black	1988-89	0.25	0.24	0.29	0.35	-0.1**	-0.1***	-0.06**
		[0.027]	[0.009]	[0.008]	[0.021]	[0.042]	[0.029]	[0.029]
	1998-99	0.26	0.25	0.29	0.35	-0.09**	-0.1***	-0.06**
		[0.029]	[0.01]	[0.007]	[0.021]	[0.043]	[0.029]	[0.026]
	2008-09	0.24	0.24	0.28	0.34	-0.09**	-0.1***	-0.06**
		[0.026]	[0.009]	[0.008]	[0.02]	[0.041]	[0.027]	[0.027]
	2018-19	0.20	0.21	0.24	0.30	-0.1***	-0.1***	-0.06**
		[0.022]	[0.007]	[0.007]	[0.018]	[0.034]	[0.023]	[0.024]
% White	1988-89	0.57	0.51	0.41	0.3	0.26***	0.21***	0.1***
		[0.027]	[0.015]	[0.005]	[0.011]	[0.028]	[0.025]	[0.014]
	1998-99	0.51	0.45	0.33	0.24	0.27***	0.21***	0.09***
		[0.028]	[0.016]	[0.007]	[0.011]	[0.028]	[0.024]	[0.014]
	2008-09	0.48	0.4	0.27	0.19	0.29***	0.21***	0.08***
		[0.027]	[0.016]	[0.007]	[0.011]	[0.029]	[0.023]	[0.014]
	2018-19	0.46	0.37	0.24	0.17	0.28***	0.2***	0.07***
		[0.024]	[0.013]	[0.006]	[0.01]	[0.028]	[0.02]	[0.013]
% Non-White	1988-89	0.43	0.49	0.59	0.7	-0.26***	-0.21***	-0.1***
		[0.027]	[0.015]	[0.005]	[0.011]	[0.028]	[0.025]	[0.014]
	1998-99	0.49	0.55	0.67	0.76	-0.27***	-0.21***	-0.09***
		[0.028]	[0.016]	[0.007]	[0.011]	[0.028]	[0.024]	[0.014]
	2008-09	0.52	0.6	0.73	0.81	-0.29***	-0.21***	-0.08***
		[0.027]	[0.016]	[0.007]	[0.011]	[0.029]	[0.023]	[0.014]
	2018-19	0.54	0.63	0.76	0.83	-0.28***	-0.2***	-0.07***
		[0.024]	[0.013]	[0.006]	[0.01]	[0.028]	[0.02]	[0.013]
Simpson's Diversity (1-D)	1988-89	0.36	0.35	0.37	0.35	<0.01	<0.01	0.02
		[0.019]	[0.006]	[0.009]	[0.017]	[0.03]	[0.02]	[0.026]
	1998-99	0.38	0.37	0.38	0.36	0.02	0.01	0.02
		[0.02]	[0.008]	[0.007]	[0.017]	[0.033]	[0.022]	[0.023]
	2008-09	0.42	0.4	0.39	0.37	0.05	0.03	0.03
		[0.023]	[0.01]	[0.005]	[0.015]	[0.036]	[0.023]	[0.017]
	2018-19	0.49	0.48	0.46	0.43	0.06*	0.05***	0.03**
		[0.021]	[0.01]	[0.006]	[0.011]	[0.03]	[0.018]	[0.013]
FE		Y	Y	Y	Y	Y	Y	Y
N		4,467	4,467	4,467	4,467	4,467	4,467	4,467

Notes: Cluster-robust SEs are in parentheses with clustering at the city-level. For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. 2018-19 total students weight all regressions. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. *** p< 0.01, ** p< 0.05, * p< 0.1.

Table 5: School Student Performance and 1935-1940 HOLC A-D Grades, Pooled 2008-2019

	Average Student Math & ELA Scores	Average Student Learning Rate (Annual)	Average Student Trend in Test Scores
A	0.643***	0.006	0.010***
	[0.04]	[0.006]	[0.003]
B	0.342***	0.008*	0.007***
	[0.028]	[0.004]	[0.002]
C	0.127***	0.005*	0.004***
	[0.015]	[0.003]	[0.001]
Constant	-0.365***	0.007***	-0.003***
	[0.013]	[0.002]	[0.001]
FE	Y	Y	Y
N	5,124	5,124	5,124

Notes: Cluster-robust standard errors are in parentheses with clustering done at the city-level. For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. All regressions are weighted by 2009-2018 total number of math and ELA tests for pooled SEDA estimates. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. *** p< 0.01, ** p< 0.05, * p< 0.1.

Online Appendix A

Table A.1: Diversity Sample Overview

Panel A: All 2018-19 U.S. Public Schools vs. HOLC Sample

	U.S. Census Bureau Region				
	Nationwide	Midwest	Northeast	South	West
Race: Percent Black	15.1%	13.6%	13.9%	22.8%	4.7%
HOLC A-D Sample	29.2%	37.2%	28.8%	40.0%	8.8%
Share	193%	274%	207%	175%	187%
Race: Percent Hispanic	27.1%	13.0%	22.0%	27.0%	42.7%
HOLC A-D Sample	37.3%	25.9%	36.4%	37.2%	57.8%
Share	138%	199%	165%	138%	135%
Race: Percent White	47.1%	64.4%	53.1%	42.0%	36.8%
HOLC A-D Sample	22.1%	27.9%	21.0%	18.1%	18.9%
Share	47%	43%	40%	43%	51%
Urban	26,070	5,333	3,791	9,092	7,854
HOLC A-D Sample	7,558	2,370	2,579	1,336	1,273
Share	29%	44%	68%	15%	16%
Suburban	30,355	6,712	7,081	9,278	7,284
HOLC A-D Sample	2,133	656	1,011	109	357
Share	7%	10%	14%	1%	5%
Charter Schools	7,340	1,473	746	2,420	2,701
HOLC A-D Sample	1,591	585	420	259	327
Share	22%	40%	56%	11%	12%
Schools	95,432	24,669	14,775	33,040	22,948
HOLC A-D Sample	9,709	3,031	3,595	1,453	1,630
Share	10%	12%	24%	4%	7%
Districts	17,741	5,913	3,640	3,826	4,362
HOLC A-D Sample	1,955	666	726	195	368
Share	11%	11%	20%	5%	8%
CBSAs	934	291	91	375	177
HOLC A-D Sample	141	51	29	46	15
Share	15%	18%	32%	12%	8%

Panel B: SEDA Sample Overview: All 2017-18 U.S. Public Districts vs. HOLC Sample

	U.S. Census Bureau Region				
	Nationwide	Midwest	Northeast	South	West
Race: Percent Black	15.2%	13.7%	13.9%	23.0%	4.8%
HOLC A-D Sample	28.9%	35.9%	28.6%	39.7%	7.7%
Share	190%	262%	206%	173%	160%
Race: Percent Hispanic	26.7%	12.7%	21.5%	26.5%	42.5%
HOLC A-D Sample	37.8%	27.9%	35.0%	38.1%	60.4%
Share	142%	220%	163%	144%	142%
Race: Percent White	47.6%	64.9%	53.9%	42.6%	37.1%
HOLC A-D Sample	22.2%	27.4%	22.3%	17.7%	17.1%
Share	47%	42%	41%	42%	46%
Urban	25,964	5,335	3,788	9,038	7,803
HOLC A-D Sample	3,847	1,432	999	799	617
Share	15%	27%	26%	9%	8%
Suburban	30,335	6,715	7,097	9,264	7,259
HOLC A-D Sample	1,272	390	590	64	228
Share	4%	6%	8%	1%	3%
Charter Schools	7,158	1,474	723	2,337	2,624
HOLC A-D Sample	505	219	155	101	30
Share	7%	15%	21%	4%	1%

Schools	95,242	24,643	14,812	32,940	22,847
HOLC A-D Sample	5,124	1,826	1,589	864	845
Share	5%	7%	11%	3%	4%
Districts	16,782	5,926	3,653	3,809	3,394
HOLC A-D Sample	1,006	393	397	127	89
Share	6%	7%	11%	3%	3%
CBSAs	934	291	91	375	177
HOLC A-D Sample	137	50	27	45	15
Share	15%	17%	30%	12%	8%

Panel C: Finance Sample Overview: All 2017-18 U.S. Public Districts vs. HOLC Sample

	U.S. Census Bureau Region				
	Nationwide	Midwest	Northeast	South	West
Race: Percent Black	15.1%	13.7%	14.1%	23.0%	4.8%
HOLC A-D Sample	24.9%	28.5%	22.4%	34.3%	7.9%
Share	165%	208%	159%	149%	165%
Race: Percent Hispanic	26.7%	12.7%	21.6%	26.5%	42.4%
HOLC A-D Sample	33.0%	20.6%	27.0%	32.1%	53.3%
Share	124%	162%	125%	121%	126%
Race: Percent White	47.6%	64.8%	53.7%	42.5%	37.2%
HOLC A-D Sample	31.7%	41.1%	39.0%	26.8%	22.3%
Share	67%	63%	73%	63%	60%
Urban	2,796	824	636	655	681
HOLC A-D Sample	822	419	190	137	76
Share	29%	51%	30%	21%	11%
Suburban	3,877	1,223	1,503	488	663
HOLC A-D Sample	893	364	404	45	80
Share	23%	30%	27%	9%	12%
Charter Schools	3,762	1,090	704	788	1,180
HOLC A-D Sample	726	368	174	108	76
Share	19%	34%	25%	14%	6%
Districts	16,799	5,924	3,668	3,812	3,395
HOLC A-D Sample	1,760	806	608	187	159
Share	10%	14%	17%	5%	5%
CBSAs	931	291	91	375	174
HOLC A-D Sample	144	53	30	46	15
Share	15%	18%	33%	12%	9%

Notes: First line of each variable includes all active 2017-18 U.S. public primary and secondary schools. Only those schools with non-zero or non-missing total student enrollment data included in sample. The second line represents all 2017-18 U.S. public primary and secondary schools for the HOLC A-D analysis sample. Schools were only included in this sample if they were matched to 1935-1940 HOLC A-D maps. Shares represent match rates for “Schools” and “CBSAs” variables and sample representativeness for all others.

Table A.2: HOLC A-D Means, SEs, and Differences,
Panel A: Full vs. Panel, Finance (USD 2018)

		Sample Comparisons			
		Full	Panel	Diff.	p-value
District PPR – Total, 2017-18 ($USD)	A	$16,938	$17,059	-$120	0.91
		[$1,778]	[$2,787]		
	B	$17,662	$17,738	-$75	0.82
		[$583]	[$906]		
	C	$15,435	$15,526	-$92	0.70
		[$399]	[$627]		
	D	$14,453	$14,269	$184	0.63
		[$472]	[$666]		
District PPR – Federal, 2017-18 ($USD)	A	$705	$695	$10	0.9
		[$133]	[$208]		
	B	$942	$932	$10	0.74
		[$52]	[$81]		
	C	$1,430	$1,411	$18	0.57
		[$49]	[$74]		
	D	$1,136	$1,088	$47	0.24
		[$52]	[$89]		
District PPR – State, 2017-18 ($USD)	A	$4,498	$4,481	$17	0.95
		[$571]	[$874]		
	B	$6,236	$6,231	$6	0.98
		[$424]	[$655]		
	C	$7,272	$7,281	-$9	0.97
		[$378]	[$593]		
	D	$7,514	$7,315	$199	0.59
		[$426]	[$536]		
District PPR – Local, 2017-18 ($USD)	A	$11,735	$11,883	-$147	0.87
		[$1,610]	[$2,535]		
	B	$10,484	$10,575	-$91	0.79
		[$604]	[$937]		
	C	$6,733	$6,834	-$101	0.57
		[$296]	[$465]		
	D	$5,803	$5,866	-$63	0.79
		[$300]	[$519]		

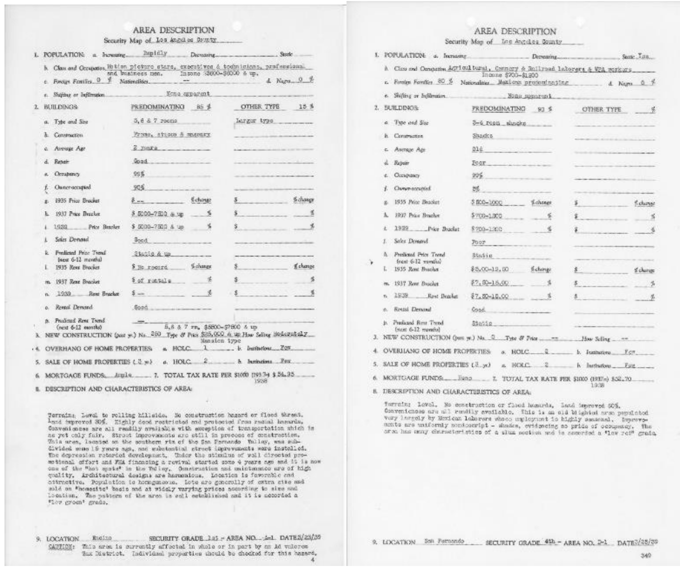
Notes: This table compares the full and longitudinal panel sample for the district finance outcomes on the different finance outcomes. Cluster-robust standard errors are in parentheses with clustering done at the city-level. No controls are included. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. All dollars denominated in USD 2018. *** p< 0.01, ** p< 0.05, * p< 0.1.

Panel B: Full vs. Panel, Racial Diversity Outcomes

		Sample Comparisons			
		Full	Panel	Diff.	p-value
% Black	A	0.22	0.20	0.018	0.047**
		[0.013]	[0.015]		
	B	0.26	0.23	0.036	<0.001***
		[0.008]	[0.009]		
	C	0.27	0.23	0.037	<0.001***
		[0.005]	[0.007]		
	D	0.36	0.30	0.056	<0.001***
		[0.007]	[0.01]		
% White	A	0.46	0.48	-0.026	0.004***
		[0.015]	[0.020]		
	B	0.33	0.37	-0.042	<0.001***
		[0.008]	[0.011]		
	C	0.21	0.25	-0.041	<0.001***
		[0.005]	[0.008]		
	D	0.13	0.15	-0.023	<0.001***
		[0.004]	[0.008]		
% Non-White	A	0.54	0.52	0.026	0.004***
		[0.015]	[0.020]		
	B	0.67	0.63	0.042	<0.001***
		[0.008]	[0.011]		
	C	0.79	0.75	0.041	<0.001***
		[0.005]	[0.008]		
	D	0.87	0.85	0.023	<0.001***
		[0.004]	[0.008]		
Simpson's	A	0.49	0.50	-0.006	0.329
Diversity (1-D)		[0.011]	[0.014]		
	B	0.480	0.50	-0.018	<0.001***
		[0.006]	[0.008]		
	C	0.45	0.45	-0.005	0.246
		[0.004]	[0.007]		
	D	0.41	0.42	-0.007	0.295
		[0.005]	[0.009]		

Notes: This table compares the full and longitudinal panel sample for the school racial diversity outcomes on the different racial diversity outcomes. Cluster-robust standard errors are in parentheses with clustering done at the city-level. No controls are included. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. *** p< 0.01, ** p< 0.05, * p< 0.1.

Figure A.1: Los Angeles County Area Descriptions for Nos. A-1 and D-1



AREA DESCRIPTION
Security Map of Los Angeles County

1. POPULATION: a. Increasing Rapidly Decreasing Static
b. Class and Occupation Motion picture stars, executives & technicians, professional and business men. Income $3000-$8000 & up.
c. Foreign Families 0 % Nationalities -- d. Negro 0 %
e. Shifting or Infiltration None apparent

2. BUILDINGS: PREDOMINATING 85 % OTHER TYPE 15 %
a. Type and Size 5,6 & 7 rooms | Larger type
b. Construction Frame, stucco & masonry
c. Average Age 2 years
d. Repair Good
e. Occupancy 99%
f. Owner-occupied 90%
g. 1935 Price Bracket $ -- % change
h. 1937 Price Bracket $ 5000-7500 & up
i. 1939 Price Bracket $ 5000-7500 & up
j. Sales Demand Good
k. Predicted Price Trend (next 6-12 months) Static & up
l. 1935 Rent Bracket $ No record
m. 1937 Rent Bracket $ of rentals
n. 1939 Rent Bracket $ --
o. Rental Demand Good
p. Predicted Rent Trend (next 6-12 months) --

3. NEW CONSTRUCTION (past yr.) No. 250 Type & Price 5,6 & 7 rm, $5800-$7800 & up; $25,000 & up Mansion type How Selling Moderately
4. OVERHANG OF HOME PROPERTIES: a. HOLC 1 b. Institutions Few
5. SALE OF HOME PROPERTIES (2 yr.) a. HOLC 2 b. Institutions Few
6. MORTGAGE FUNDS: Ample 7. TOTAL TAX RATE PER $1000 (1937-1938) $ 54.35
8. DESCRIPTION AND CHARACTERISTICS OF AREA:

Terrain: Level to rolling hillside. No construction hazard or flood threat. Land improved 30%. Highly deed restricted and protected from racial hazards. Conveniences are all readily available with exception of transportation which is as yet only fair. Street improvements are still in process of construction. This area, located on the southern rim of the San Fernando Valley, was subdivided some 15 years ago, and substantial street improvements were installed. The depression retarded development. Under the stimulus of well directed promotional effort and FHA financing a revival started some 4 years ago and it is now one of the "hot spots" in the Valley. Construction and maintenance are of high quality. Architectural designs are harmonious. Location is favorable and attractive. Population is homogeneous. Lots are generally of extra size and sold on "homesite" basis and at widely varying prices according to size and location. The pattern of the area is well established and it is accorded a "low green" grade.

9. LOCATION Encino SECURITY GRADE 1st AREA NO. A-1 DATE 3/23/39
CAUTION: This area is currently affected in whole or in part by an Ad valorem Tax District. Individual properties should be checked for this hazard.
4

AREA DESCRIPTION
Security Map of Los Angeles County

1. POPULATION: a. Increasing Decreasing Static Yes
b. Class and Occupation Agricultural, Cannery & Railroad laborers & WPA workers Income $700-$1200
c. Foreign Families 80 % Nationalities Mexican predominating d. Negro 0 %
e. Shifting or Infiltration None apparent

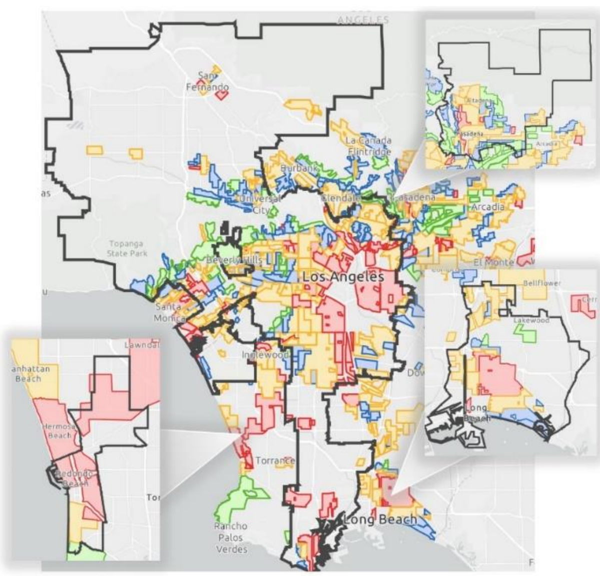
2. BUILDINGS: PREDOMINATING 90 % OTHER TYPE %
a. Type and Size 3-4 room shacks
b. Construction Shacks
c. Average Age Old
d. Repair Poor
e. Occupancy 99%
f. Owner-occupied 5%
g. 1935 Price Bracket $ 800-1000 % change
h. 1937 Price Bracket $700-1200
i. 1939 Price Bracket $700-1200
j. Sales Demand Poor
k. Predicted Price Trend (next 6-12 months) Static
l. 1935 Rent Bracket $5.00-12.50 % change
m. 1937 Rent Bracket $7.50-15.00
n. 1939 Rent Bracket $7.50-15.00
o. Rental Demand Good
p. Predicted Rent Trend (next 6-12 months) Static

3. NEW CONSTRUCTION (past yr.) No. 0 Type & Price -- How Selling --
4. OVERHANG OF HOME PROPERTIES: a. HOLC 2 b. Institutions Few
5. SALE OF HOME PROPERTIES (3 yr.) a. HOLC 2 b. Institutions Few
6. MORTGAGE FUNDS: None 7. TOTAL TAX RATE PER $1000 (1937-1938) $52.70
8. DESCRIPTION AND CHARACTERISTICS OF AREA:

Terrain: Level. No construction or flood hazards. Land improved 50%. Conveniences are all readily available. This is an old blighted area populated very largely by Mexican laborers whose employment is highly seasonal. Improvements are uniformly nondescript - shacks, evidencing no pride of occupancy. The area has many characteristics of a slum section and is accorded a "low red" grade.

9. LOCATION San Fernando SECURITY GRADE 4th AREA NO. D-1 DATE 3/28/39
349

Figure A.2: Los Angeles and Surrounding Districts & 1935 – 1940 HOLC Maps



Notes: "Best" (A, outlined in green), "Still Desirable" (B, outlined in blue), "Definitely Declining" (C, outlined in yellow), to "Hazardous" (D, outlined in red) - [bottom left] Redondo Beach Unified; [center] Los Angeles Unified; [top right] Pasadena ISD; [bottom right] Long Beach Unified

Figure A.3: School Total PPE by HOLC A-D District Grades – Only "D" Schools (USD 2018)

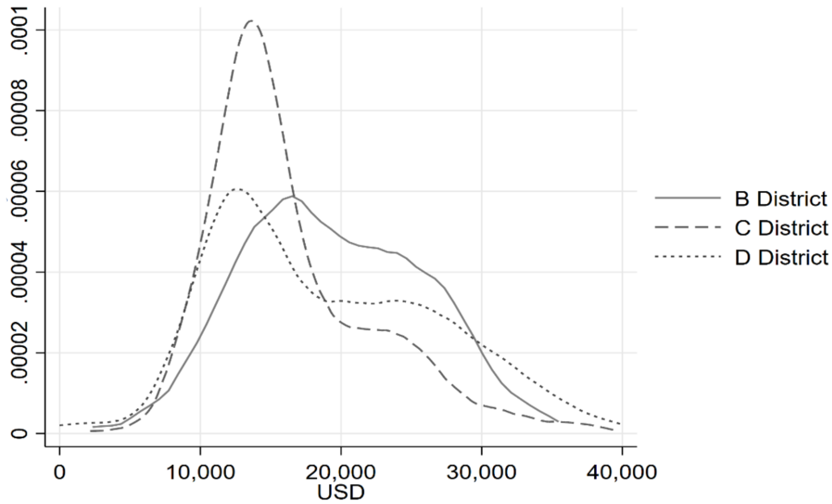
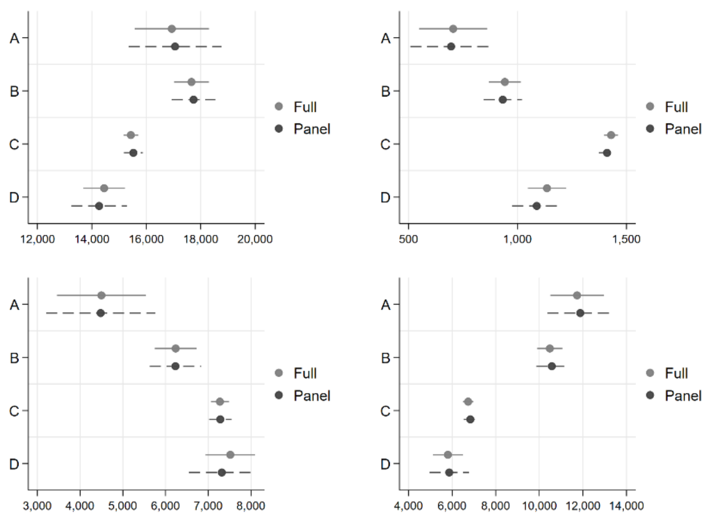
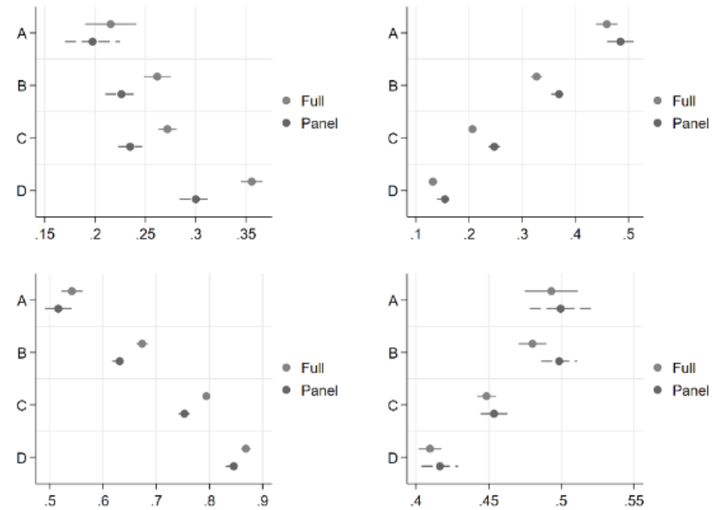


Figure A.4: HOLC A-D Coefficient Plots, Full vs. Panel, 2017-18, Finance (USD 2018)



Notes: [top left] Per-Pupil Total Revenue; [top right] Per-Pupil Federal Revenue; [bottom left] Per-Pupil State Revenue; [bottom right] Per-Pupil Local Revenue. All values represented in USD 2018. Coefficients (solid dots) are HOLC A-D weighted averages from regressions of respective finance outcomes on HOLC A-D indicators without a constant term. The 95% confidence intervals (lines) are calculated using Huber-White heteroskedasticity-consistent standard errors. Regressions are run using both the 2017-18 full sample and the 2017-18 data from the panel sample. The full sample is the original 2017-18 cross-sectional sample and includes all d = 1,760 districts. The panel sample is the time series sample spanning 1989-90 through 2017-18 school years and includes d_p = 1,109 districts.

Figure A.5: HOLC A-D Coefficient Plots, Full vs. Panel, Diversity Outcomes



Notes: [top left] Percent Black; [top right] Percent White; [bottom left] %Non-White; [bottom right] Simpson's Diversity (1-D). Comments on full and panel samples here. Coefficients (solid dots) are HOLC A-D weighted averages from regressions of respective diversity outcomes on HOLC A-D indicators without a constant term. The 95% confidence intervals (lines) are calculated using Huber-White heteroskedasticity-consistent standard errors. Regressions are run using both the 2017-18 full sample and the 2017-18 data from the panel sample. The full sample is the original 2017-18 cross-sectional sample and includes all n = 9,709 schools. The panel sample is the time series sample spanning 1988-89 through 2018-19 school years and includes np = 4,677 schools.

Online Appendix B

NCES Geospatial Data

We combine the HOLC data with geospatial district and school-level data from the National Center for Education Statistics (NCES). Our geospatial data includes district boundaries and public school-level latitude and longitude point locations. The 2018-2019 district boundary data derives from the Census TIGER/2019 Line geospatial data (U.S. Census Bureau, 2020).[23] We use the 2018-19 NCES EDGE school-point location data to complete our school-level analyses, which provide latitude and longitude coordinates for public elementary and secondary schools from the NCES EDGE Common Core of Data (CCD) (U.S. Department of Education, 2020).[24]

NCES Non-Geospatial Data

The NCES non-geospatial data we use is also at the district and school levels. This data includes fiscal and non-fiscal data, which we leverage in our analyses below on district-level financing (i.e., local, state, federal, total) and school-level racial diversity by HOLC grade. We leveraged the most recent NCES 2017-18 F-33 survey data for the fiscal data, which provides general financing information (e.g., revenue and expenditure totals and subtotals) at the district level. This data is provided through NCES CCD (U.S. Department of Education, 2020). We use F-33 reported total general revenues for district-level revenues and their associated first-level revenue subtotals (e.g., local, state, federal). We transform all district-level financial data into per-pupil terms to account for differences in enrollment.[25]

For the non-fiscal district and school data, we use NCES 2017-2018 and NCES 2018-19 datasets. We analyze each school's Simpson's Diversity Index (1-D) for our within-school-between-student racial diversity outcome. This index captures the likelihood that two randomly selected students from a given school will belong to different racial groups, ranging from 0 to 1, with larger values representing greater within-school-between-student racial diversity (Simpson, 1949; Hirschman, 1964).[26]

National Education Resource Database on Schools Data

To complement our district-level finance analysis, we combine the HOLC data with school-level finance data from the Edunomics Lab at Georgetown University (Edunomics Lab, 2021). Their newly released National Education Resource Database on Schools (NERD$) data includes state-reported 2018-19 school-level expenditures across 49 states and the District of Columbia and captures per-pupil total expenditures, federal expenditures, and combined state

[23] We supplement this data with 2018-2019 NYC public school district boundaries from the NYC Department of City Planning (DCP). This supplemental geospatial boundary data was required since all NYC public school districts were reported as one unified district in the NCES EDGE geospatial data. Thus, combining the NYC DCP and NCES EDGE geospatial data allow us to account for each of NYCDOE's 32 school districts and more accurately capture the within-district variation in HOLC grades across the city (NYC Department of City Planning, 2020).

[24] Geospatial school boundary files also exist in the NCES School Attendance Boundary Survey (SABS). SABS was an experimental survey led by NCES and supported by the U.S. Census Bureau to collect school boundaries for the 2013-14 and 2015-16 school years. This effort led to the collection of over 70,000 school boundaries across the United States but is now discontinued (U.S. Department of Education, 2020). Due to incomplete nationwide coverage, we opt not to use SABS in this study.

[25] As a result, we use district enrollment weights to calculate per-pupil weighted averages by HOLC A-D security rating. In addition, our focus on district revenues versus district expenditures is an artifact of variable coverage. NCES district expenditures are not disaggregated at the local, state, and federal levels, but NCES district revenues are. When comparisons are possible, district expenditure results (i.e., per-pupil total, instructional salaries, benefits) mirror district revenues (i.e., per-pupil total).

[26] The equation is $1 - D_i = 1 - \sum_{r=1}^{R} p_r^2$, where p_r represents the probability that two randomly selected students from a given i^{th} school will belong to the same race, and *r* represents the seven commonly used racial categories by the U.S. Census Bureau, including American Indian or Alaska Native, Asian, Black or African American, Hispanic, Native Hawaiian or Other Pacific Islander, White, and Multiple Races.

and local expenditures.[27] Each per-pupil outcome we use in our analysis was provided directly by NERD$ and based on state-reported school-level student enrollment versus NCES school-level student enrollment. As a result, there are some discrepancies between the state reported and NCES school enrollment data, due to varying enrollment metrics (e.g., September Count Day, October County Day, Average Daily Attendance (ADA), Average Daily Memberships (ADM), Weighted Enrollment) used by states. However, these differences are often immaterial and do not change our overall school-level findings.

Stanford Education Data Archive Data

We used school-level Stanford Education Data Archive (SEDA) data for all our school-performance analyses. This data provides students' academic outcomes in grades 3-8, spanning SY2008-09 to SY2017-18, and includes students' average test scores, test score trends, and learning rates. Average test scores are a school's mean test-based achievement pooled across Math and ELA subjects. Learning rates are the school-grade slope of a school's mean test-based achievement or how test scores change across grades within a cohort. Test score trends are the school-cohort slope of a school's mean test-based achievement or how test scores change across student cohorts within grades (Reardon et al., 2021). Like SEDA, we use "educational opportunities" and "changes in educational opportunities" interchangeably with "average test scores" and "test score trends," respectively. The terms are substituted for another in the sense that test score changes reflect educational opportunities in a community and are influenced by occasions to learn at home, in neighborhoods, in child-care centers, at preschool and after-school programs, and from peers at school (Reardon et al., 2021). We rely on SEDA's cohort standardized (C.S.) scale achievement estimates based on the Spring 2009 4th grade cohort for each student achievement measure. The C.S. scale achievement estimates are measured in S.D. units relative to the national average and are calculated using OLS and Empirical Bayes (Reardon et al., 2021).[28] Pairing these measures with 1935-1940 HOLC maps allows us to look at how HOLC redlining maps from over eight decades ago are associated with the current state of educational opportunity and student learning rates today.

[27] By law, districts must report actual dollars spent instead of estimations based on teacher FTE counts and average teacher salaries. From my discussions with NERD$, nearly every district says they report actual dollars; however, NERD$ occasionally notices within-district trends that suggest otherwise. Thus, overall NERD$ data will be actual dollars spent, but sometimes it will be based on teacher FTE and average teacher salaries.

[28] As per SEDA, OLS estimates are more appropriate here than E.B. estimates since we use precision weights in our regression models. Regardless, our results are robust to the underlying estimation procedure.

Online Appendix C

Table C.1 School-Level Outcomes with Covariates

Nationwide	School PPE Total	School PPE Federal	School PPE State & Local	% Black	% White	% Non-White	Simpson's Diversity	Av. Math & ELA Scores	Avg. Learning Rate (Annual)	Avg. Trend in Test Scores
A	-2909.48	-589.489***	-2308.13	-0.106***	0.262***	-0.262***	0.049	0.505***	0.002	0.007
	[1926.448]	[51.253]	[1927.831]	[0.037]	[0.027]	[0.027]	[0.039]	[0.051]	[0.010]	[0.006]
B	-1971.77	-318.777***	-1654.44	-0.052**	0.161***	-0.161***	0.033	0.303***	0.002	0.004
	[1263.503]	[39.466]	[1247.270]	[0.024]	[0.016]	[0.016]	[0.027]	[0.034]	[0.004]	[0.004]
C	-1343.780*	-131.028***	-1209.437*	-0.022	0.035***	-0.035***	0.015	0.091***	0.002	0.001
	[706.938]	[40.666]	[678.076]	[0.020]	[0.006]	[0.006]	[0.015]	[0.014]	[0.003]	[0.002]
Family Income	44.042*	-6.497**	50.138**	0.003*	0.002	-0.002	0.001	0.011**	0.001	0.000***
	[25.288]	[2.658]	[24.108]	[0.002]	[0.001]	[0.001]	[0.001]	[0.004]	[0.000]	[0.000]
Building Age	-15.538	0.942	-16.773	0	-0.001	-0.001	0	0	0	0
	[18.744]	[1.438]	[17.932]	[0.001]	[0.001]	[0.001]	[0.000]	[0.001]	[0.000]	[0.000]
% Black	-5.968	0.239	-6.099	0.004***	-0.001***	-0.001***	-0.001***	-0.002*	0	-0.000*
	[7.481]	[0.749]	[7.085]	[0.001]	[0.000]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]
Constant	18535.11	1372.005	17160.9	0.294	0.162	0.838	0.43	-0.357	0.009	-0.006
	[1053.728]	[53.960]	[1020.116]	[0.029]	[0.018]	[0.018]	[0.017]	[0.033]	[0.005]	[0.004]
N	6553	6553	6553	6553	6553	6553	6553	3906	3906	3906

Notes: Cluster-robust standard errors are in parentheses with clustering done at the city-level. For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. Controls for family income, percent Black, and building age are included. 2017-18 total students weight all regressions. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. All dollars denominated in USD 2018. *** p< 0.01, ** p< 0.05, * p< 0.1.