ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 21-363

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

Dylan Lukes
Harvard University

Christopher Cleveland
Harvard University

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, which arguably influenced banks and other 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 school and district funding, school racial diversity, and school performance. We find that schools and districts located today in historically redlined D neighborhoods have less district per-pupil total revenues, larger shares of Black and non-White student bodies, less diverse student populations, and worse average test scores relative to those located in A, B, and C neighborhoods. Conversely, at the school level, we find that per-pupil total expenditures are better for those schools operating in previously redlined D neighborhoods. Consequently, these schools also have the largest shares of low-income students. Our nationwide results are, on the whole, consistent by region and after controlling for CBSA. Finally, we document a persistence in these patterns across time, with overall positive time trends regardless of HOLC security rating but widening gaps between D vs. A, B, and C outcomes. These findings suggest that education policymakers need to consider the historical implications of redlining and past neighborhood inequality on neighborhoods today when designing modern interventions focused on improving the life outcomes of students of color and students from low-socioeconomic backgrounds.

VERSION: November 2021

Suggested citation: Lukes, Dylan, and Christopher Cleveland. (2021). The Lingering Legacy of Redlining on School Funding, Diversity, and Performance. (EdWorkingPaper: 21-363). Retrieved from Annenberg Institute at Brown University: https://-doi.org/10.26300/qeer-8c25

A Lingering Legacy: The Relationship Between 1930s HOLC "Redlining" Maps and School Funding, Diversity, and Performance

Dylan Lukes[a,*]

Harvard University

Christopher Cleveland[a]

Harvard University

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, which arguably influenced banks and other 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 school and district funding, school racial diversity, and school performance. We find that schools and districts located today in historically redlined D neighborhoods have less district per-pupil total revenues, larger shares of Black and non-White student bodies, less diverse student populations, and worse average test scores relative to those located in A, B, and C neighborhoods. Conversely, at the school level, we find that per-pupil total expenditures are better for those schools operating in previously redlined D neighborhoods. Consequently, these schools also have the largest shares of low-income students. Our nationwide results are, on the whole, consistent by region and after controlling for CBSA. Finally, we document a persistence in these patterns across time, with overall positive time trends regardless of HOLC security rating but widening gaps between D vs. A, B, and C outcomes. These findings suggest that education policymakers need to consider the historical implications of redlining and past neighborhood inequality on neighborhoods today when designing modern interventions focused on improving the life outcomes of students of color and students from low-socioeconomic backgrounds.

Key Words: HOLC, Redlining, Educational Inequality, School Finance, School Diversity, School Performance

JEL Codes: I20, I21, I22, I24

---

[*] Corresponding author.
[a] Harvard Graduate School of Education, 13 Appian Way, Cambridge, MA 02138; Email addresses: dylanlukes@g.harvard.edu (D. Lukes), chcleveland@g.harvard.edu (C. Cleveland).

1. INTRO

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.[1] 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.,

[1] HOLC and Federal Housing Administration (FHA) maps are not one and 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, also under the FHLBB, created their own 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 a few important findings. First for district-level funding, we find present-day districts located in predominantly D neighborhoods have less district-level per-pupil total revenues, but generally higher per-pupil federal and state revenues than those districts 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 we observe 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 likely 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 relatively 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 virtually no differences in either measures of average learning rates or educational opportunity changes by A-D HOLC grade. These results are generally 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 2009-2018.

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 both 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 organized as follows. Section II 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 III describes the data used for this research. Section IV/V outlines this paper's methods, including how analysis samples were constructed and details on the analytic approach. Section VI provides the results. Section VII discusses and 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 5 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. As an 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 negro 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* captured 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.[2]

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,

[2] The area descriptions for Los Angeles A-1 and D-1 can be viewed in Figure A1 in Appendix A.

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 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 the modern demographic and residential patterns of HOLC-graded areas. 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 that currently have relatively 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 regression discontinuity 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 similar to Aaronson et al. (2020), they find sizeable and statistically significant differences across all outcomes that favor higher-rated HOLC neighborhoods to 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 time of HOLC-map creation and find 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 not consistent 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 nearly all areas in their study, previous HOLC redlined areas have consistently elevated land surface temperatures compared to non-redlined areas. Regionally, 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 the levels of federal, state, local, and within-district. The two primary federal revenue sources are Title I and IDEA that generally 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 generally 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 greatly value school spending including spending on teachers and staff salaries. Additionally, they show that salary spending is allocated inefficiently throughout much of 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 economically disadvantaged 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 several 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 blacks in the United States to the unparalleled degree of deliberate segregation they experience in American cities. Chetty, Hendren, and Katz (2016) find 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 district and school-level data sources we used in our analysis, in addition to the digitized Home Owners' Loan Corporation (HOLC) maps created by the "Mapping Inequality" project (Nelson et al., 2020). For further details please refer to Appendix Tables A1-A4.

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 1935-1940 in Shapefile or GeoJSON format. The spatial data includes over 7,000 neighborhoods and nearly 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). 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. 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 both district boundaries and public school-level latitude and longitude point locations. The 2018-2019 district boundary data used derives from the Census TIGER/2019 Line geospatial data (U.S. Census Bureau, 2020).[3] To complete our school-level analyses, we use the 2018-19 NCES EDGE school-point location data, which provides 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).[4]

[3] 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).
[4] Geospatial school boundary files also exist as part of 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.

3.3. NCES Non-Geospatial Data

The NCES non-geospatial data we use is also at the district and school level. This data includes both 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. For the fiscal data, we leveraged 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. This data is provided through NCES CCD (U.S. Department of Education, 2020). For district-level revenues, we use F-33 reported total general 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.[5]

For the non-fiscal district and school data, we use NCES 2017-2018 and NCES 2018-19 datasets. For our within-school-between-student racial diversity outcome, we analyze each school's Simpson's Diversity Index (1-D). This index captures the likelihood that two randomly selected students from a given i^{th} school will belong to different racial groups, and ranges from 0 to 1 with larger values representing greater within-school-between-student racial diversity (Simpson, 1949; Hirschman, 1964).[6] As a complement to Simpson's Diversity Index, we also consider the Exposure Index, a popular segregation index in the social sciences and one that approximates between-group contact and interaction for a given race-group pair (Massey & Denton, 1988). We use this index to evaluate segregation for all schools in our sample by A-D

[5] 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 largely 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 those of district revenues (i.e., per-pupil total).

[6] 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. About three percent of the schools in our sample were missing student body breakdowns by race due to data quality flags or failure to report. We omit these schools from our analysis.

HOLC grade for various majority-minority race-group pairs, including White-Asian, White-Black, White-Hispanic and White-Non-White.

3.4. National Education Resource Database on Schools (NERD$) Data

To complement our district-level finance analysis, we also 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$) data includes state-reported 2018-19 school-level expenditures across 49 states and the District of Columbia and captures per-pupil total expenditures, federal expenditures, and combined state and local expenditures. Each per-pupil outcome we use in our analysis was provided directly by NERD$ 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, likely 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.

3.5. Stanford Education Data Archive (SEDA) Data

We used school-level Stanford Education Data Archive (SEDA) data for all our school-performance analyses. This data provides us with students' academic outcomes grades 3-8 spanning 2009-18 and includes students' average test scores, test score trends, and learning rates. "Average test scores" and "test score trends" reflect educational opportunities and changes therein, while "learning rates" measures the amount learned per grade (Reardon et al., 2021). We rely on SEDA's cohort standardized (CS) scale achievement estimates based on the Spring 2009 4th grade cohort for each of these student achievement measures. The CS scale achievement estimates are measured in SD units relative to the national average and are calculated using OLS and Empirical Bayes (Reardon et al., 2021).[7] 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.

4. ANALYSIS SAMPLES

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.[8] In what follows, we provide an overview of how we constructed each school and district-level analysis sample, and describe our analytic approach.

4.1. School-Level Analysis Samples

A key piece of this paper is the connection between the geospatial HOLC maps and the geospatial NCES public district and school-level data. For the NCES school-level geospatial data, our mapping approach is straightforward. If a modern-day school, based on its latitude and longitude point location, is contained within a historic HOLC A-D polygon, it is assigned that historic HOLC A-D polygon grade. If not, no grade is assigned, and the school is dropped from

[7] As per SEDA, OLS estimates are more appropriate here than EB estimates since we use precision weights in our regression models. Regardless, our results are robust to the underlying estimation procedure.
[8] 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.

the sample. Thus, a school is assigned, at most, one HOLC A-D grade.[9] This approach, applied to all 2018-19 U.S. public primary and secondary schools, links s = 10,013 public schools to a unique HOLC A-D HOLC grade. From here, we merge this data with the NCES 2018-19 school-level data to create our primary cross-sectional "Student Racial Diversity" analysis sample. After accounting for missing data, the "Student Racial Diversity" analysis sample has s = 9,709 U.S. public primary and secondary schools in d = 1,955 unique districts (Table 1).[10] This represents just over 10% of all active 2018-2019 U.S. public primary and secondary schools. We also link these schools to their late 1980s, 1990s, and 2000s school demographics for our subsequent longitudinal analysis. Doing so creates a "Student Racial Diversity" longitudinal analysis sample with s = 4,677 schools in d = 590 districts covering 1989-2019 (Table 1).

To construct the school-level "Student Performance" analysis sample, we follow the same exercise described above, but instead of linking the HOLC A-D dataset to NCES 2018-19 demographic data, we now link it to SEDA 2009-18 student performance data. Doing so reduces the sample for the school-level performance analysis by just over one quarter, from s = 10,013 to s = 7,303 U.S. public primary and secondary schools mapped to unique HOLC A-D grades. Nearly all unmatched schools in the SEDA dataset to the HOLC A-D dataset are due to grade-level mismatches (i.e., SEDA data only covers grades 3-8). Some unmatched schools are also due to school year coverage (i.e., SEDA data only covers schools up to 2018). As a result, all grade levels K-2 and 9-12 and any new public 3-8 grade level schools created after 2018 are

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

[10] The initial merge with 2018-19 NCES school-level data matches 98.5% of schools. The remaining unmatched 1.5% of schools are subsequently matched with NCES 2017-18 data. Only 1 school remains after matching 2018-19 orphans with 2017-18 NCES schools. This school is dropped from the data giving s = 10,012. Finally, two schools mapped to the E HOLC rating, where E represents "other", were also dropped giving us a total of s = 10,010 observations for the school-level diversity analysis sample. After accounting for missing school demographic data across our outcome variables (i.e., Percent Black, White, Non-White, Simpson's Diversity Index) we are left with a total of s = 9,709.

excluded from this sample. Finally, after removing matched schools with missing SEDA student performance data, the sample drops to s = 5,124 schools across d = 1,006 districts. The "Student Performance" analysis sample, except for a small school set due to missing student enrollment data, perfectly overlaps with the more extensive "Student Racial Diversity" sample.[11]

While the "Student Performance" and "Student Racial Diversity" analysis samples differ notably in size, on average, they share similar student demographics and other school characteristics. For example, in each sample, Black students, on average, make up about 30 percent of the student population, double that of the national average for all U.S. public schools. Hispanic students are also consistently oversampled in both "Student Performance" and "Student Racial Diversity" samples relative to the U.S. national average, making up around 37 percent of the student population relative to 27 percent, respectively. In contrast, White students are underrepresented, making up just over 22 percent of student populations in both samples. At the regional level, the samples have very similar Black, Hispanic, and White splits across each U.S. Census Bureau region, with once again consistent oversampling of Black and Hispanic students and under-sampling of White students (Table A2-A3). Unsurprisingly, since HOLC A-D maps predominantly targeted larger urban areas, schools in both samples are also almost exclusively located in urban (75 percent) or suburban (20 percent) locales, with less than 5 percent in rural areas or towns. On average, roughly 55 percent of all U.S. public schools are in urban (27 percent) and suburban locales (32 percent), indicating our HOLC A-D samples contain a disproportionate number of these locales relative to national averages (Table A2-A3). Finally, while the "Student Racial Diversity" sample captures a few additional CBSAs compared to the

[11] The "Student Performance" sample has 5 districts and 7 schools that the "Student Diversity" sample doesn't. This is due solely to schools in the "Student Diversity" sample being dropped as a result of missing student enrollment and demographic data, not because of failed NCES school ID matches.

"Student Performance" sample, both have very similar coverage across all U.S. cities included in the sample.[12] For a more thorough overview of the school-level analysis samples, please refer to Appendix A, where we have included sample breakdowns by CBSA and HOLC security ratings and an overview of sample characteristics with comparisons to national averages for all U.S. public schools (Table A1-A4).

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 in some cases, entire counties, and as a result, at times envelope multiple HOLC A-D polygons. Thus, while for schools we could use a one-to-one mapping procedure based on school point locations, for districts, we use a one-to-many. To do so, 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 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 which 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. From here, we use this dataset to calculate the HOLC A-D weighted average for each public school district, with weights based on HOLC A-D polygon areas.

[12] Importantly, the distribution of HOLC A-D grades remains stable across both school-level diversity and SEDA samples, both nationwide and by CBSA (Table A4). HOLC A-D distributions are nearly identical at the national level, varying only by a few percentage points. Comparing samples by CBSA, we see almost the same set of CBSAs represented across the school-level diversity and SEDA samples. In addition, while there is a bit more variation in HOLC A-D distributions across samples at the CBSA level, these are generally driven by CBSAs with already low district counts in the "School Diversity" analysis sample.

After this mapping exercise is complete, each public school district included in the NCES 2018-19 geospatial data now has a HOLC A-D grade assigned to it; however, not all public-school districts have polygon boundaries (e.g., charter schools). To capture these omitted districts, we use NCES 2018-19 school district geospatial data with district latitude and longitude point locations. Following a similar procedure outlined above for our school-level point location mapping to HOLC A-D grades, we obtain HOLC A-D grades for each previously omitted district point location. Next, we append the polygon district datasets and point location district datasets, each with their respective HOLC grade mappings, to create our district to HOLC grade mapping dataset.[13] After merging 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 of d = 1,760 unique districts, which captures just over 10% of all U.S. public school districts (Table 1, Table A1). Nearly all unmatched NCES districts between the HOLC grade mappings dataset and 2017-18 F33 finance data are due to a mismatch in school year coverage (i.e., F-33 data only covers the 2017-18 school year).

Like the school-level analysis samples described above, the "Fiscal" analysis sample oversamples districts from urban and suburban areas, and those that educate larger shares of Black and Hispanic students relative to the U.S. national average. This is once again true both overall and by U.S. Census Bureau region. Overlap between the "Fiscal" and school-level analysis samples described above is considerable. Around 80 percent of schools in the "Student Racial Diversity" analysis sample are captured by those districts in the "Fiscal" sample, and just over 90 percent of schools in the "Student Performance" analysis sample. As before, we provide a more thorough overview of analysis samples in the Appendix (Tables A1-A4). In addition, to

[13] Given the natural overlap between the NCES 2018-19 district polygon boundary dataset and the NCES 2018-19 district point location dataset, duplicates from the point location dataset were removed once appended.

view examples of HOLC A-D maps overlaid with NCES district boundaries, please refer to Appendix Figures A2 and A3 where we provide NCES district boundary maps superimposed on 1935-1940 HOLC maps for a small selection of CBSAs in our samples (i.e., Cleveland Municipal School District, Los Angeles Unified School District).

[Insert Table 1 here]

5. EMPIRICAL STRATEGY

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.[14] 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.[15] 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 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:

[14] 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.

[15] Recall, for districts, this is not quite 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.

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

where i indexes schools or districts, j indexes CBSAs, and γ_j is the CBSA fixed effect. For schools, $HOLC_A$, $HOLC_B$ and $HOLC_C$ are the estimated average differences in the outcome Y_{ij} between schools located today in what once were redlined HOLC D neighborhoods and their contemporaries 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 historically assigned HOLC A-D grades, and in many cases, did not even exist yet, selection bias remains a threat if past HOLC A-D neighborhood grade assignments captured pre-existing inequality between neighborhoods which in turn is a determinant of our modern-day educational outcomes. For example, suppose the original 1935-1940 HOLC A-D grade assignment reflected already existing neighborhood variation (e.g., homeownership rates, home values, access to public services, crime, etc.) that favored higher-rated HOLC security ratings (i.e., A, B) relative to lower ones (i.e., C, D). In this case, these differences may be driving, in part, the differences in educational outcomes by HOLC A-D grades we see today. If true, our results are not solely attributable to the HOLC A-D security ratings themselves but differences in other pre-existing neighborhood characteristics. Notably, Aaronson et al. (2020) find evidence that this is indeed the case while also subsequently presenting causal evidence in Aaronson et al. (2021). Given these concerns, we are careful to avoid framing our results as causal and instead focus on describing the historical long-term correlational relationships between HOLC A-D grades and their conceptions of neighborhood quality and modern-day educational outcomes.

6. RESULTS

To facilitate a more meaningful discussion of our results, we partition them by the three overarching outcomes addressed in this paper: (1) district and school-level financing, (2) school-level racial diversity, and (3) school-level student performance. In each section, we provide and discuss the regression results from our CBSA FE estimation model described in the "Empirical Strategy" section above. For each outcome, we also provide a time series analysis that looks at how the relationship between HOLC A-D grades and respective outcomes has changed over the past three decades that includes district-level financing and school-level racial diversity covariates.[16] A discussion of our results are reserved for the final section of the paper.

6.1. District & School Finance

6.1.1. Current Outcomes – Districts

In Table 2, we see that both nationwide and by region, average total per-pupil revenues are almost always lower for districts mapped to HOLC D grades than those mapped to A, B, and

[16] 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.

C grades. Nationwide, districts mapped to HOLC D grades have, on average, $14,402 per-pupil total revenues with differences greatest between A vs. D districts ($1,546, SE = $822) and least between B vs. D districts ($992, SE = $536). All of these results are statistically significant. Across regions, while point estimates for HOLC A-C indicators are almost always positive, many lack statistical significance. This result is, at least in part, a consequence of reduced statistical power given smaller region-specific subsamples. Of those HOLC A-C indicators that remain statistically significant in each region, all but the B vs. D difference in the West region align with the nationwide narrative, albeit with considerable regional heterogeneity in baseline per-pupil total revenues. For example, the Northeast has the highest average per-pupil total revenues across all HOLC grades. In contrast, the South exhibits some of the lowest per-pupil total revenue averages in our sample.

[Insert Table 2 here]

Nationwide, districts mapped to HOLC D grades have, on average, $1,201 in per-pupil federal revenues with A (-$722, SE = $139) and B (-$391, SE = $87) districts being outstripped in federal funding by D districts, and C districts ($253, SE = $76) just edging out D districts. All of these differences are statistically significant and are relatively uniform across regions. At the state level, D districts have, on average, $7,960 in per-pupil state funding with differences largest between A vs. D districts ($3,670, SE = $752) and smallest between C vs. D districts (-$461, SE = $495). All but the C vs. D average comparisons are statistically significant. While magnitudes in baseline state funding levels differ noticeably across regions, with again the Northeast leading and the South trailing, we see similar regional patterns in D vs. A, B, C differences as we observe nationwide. Apart from C vs. D differences, all are statistically significant.

Finally, at the local level, we see that nationwide 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. Regionally, we see similar patterns play out with average local funding levels monotonically decreasing as one moves from the highest (e.g., A) to lowest (e.g., D) HOLC grade. Apart from two C vs. D differences in the Midwest and South, all gaps in per-pupil local funding between D vs. A, B, and C districts are statistically significant. By combining these results with our above state, federal and total results, a cohesive narrative emerges. Districts mapped to HOLC D grades have on average significantly less per-pupil local funding than higher-rated HOLC A, B, and C districts in the sample. This gap in local funding is reduced by 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 in local district financing.

6.1.2. 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, SE = $852) and smallest between schools mapped to C and D HOLC grades (-$882, SE = $334). These differences are all statistically significant. Looking across regions, we see nearly identical expenditure patterns by HOLC grade, albeit with level shifts in baseline per-pupil expenditures up (i.e., Northeast) and down (i.e., Midwest, South, West) which comport with nationally representative statistics in per-pupil school expenditures by U.S. region (U.S. Census, 2021). The Midwest stands out as the lone exception to this rule, where point estimates are positive and null for each HOLC regression indicator, indicating equality in per-pupil expenditures regardless of past HOLC A-D neighborhood security ratings for this region.

While the 2018-19 NERDS data does afford us access to school-level finance data, it is also more coarsely disaggregated relative to the NCES F-33 district finance data, limiting our ability to deconstruct the underlying variation in 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 (Table 3). 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, SE = $39) and smallest between A and C schools ($141, SE = $31).

[Insert Table 3 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, SE = $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.

6.1.3. Reconciliation of Results

We reconcile the differences in the district and school-level results with a few explanations. One explanation, differences in measurement. Our district-level analysis uses per-pupil revenues, whereas our school-level analysis uses per-pupil expenditures.[17] If notable differences exist in allocations versus spending, this could be driving some of the differences we observe. That said, we do not see notable differences in our district-level results when we substitute per-pupil total expenditures for per-pupil total revenues.

Another explanation for the differences we observe in the district and school-level results is the school and district mapping strategies we use to assign HOLC grades. Recall, for HOLC districts, mapping is based on the weighted average of all HOLC A-D polygons contained within that district; however, for HOLC schools, 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

[17] 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.

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

[Insert Figure 2 here]

To check this, we construct the underlying HOLC A-D school distributions of each HOLC A-D district. In Figure 2, a few notable patterns emerge. First, higher (lower) rated HOLC districts have larger shares of higher (lower) rated HOLC schools, with the greatest being that of 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, too. These results validate our school and district mapping strategies and suggest schools are not clustering in polygons discordant from the mapped district HOLC A-D grade.

A final explanation, and one we believe is most likely, is that districts are simply allocating resources to those most in need (e.g., low-income students), which are those schools located today in historically 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, SE = 0.02) and smallest between A vs. C schools (-0.06, SE = 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 districts, regardless of their A-D assigned grade, systematically target and allocate more money to schools that happen to be located in historically HOLC D neighborhoods because those schools 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 provides some of the first evidence of secondary ripple effects that stem from the adverse impacts HOLC maps had on neighborhood quality and development.

Finally, considering our school-level findings, what level of importance should one place on the district-level results? We argue these 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. Figure 3 illustrates this point, where kernel densities of per-pupil total funding for D schools are shifted to the right for districts mapped to higher HOLC grades. We see a similar distributional pattern for C schools, too. These distributional patterns suggest that districts with more financial resources (i.e., A, B) allocate more to D schools in their district than those districts with fewer financial resources (i.e., C, D). All distributional differences are statistically significant.

[Insert Figure 3 here]

To confirm this, we regress per-pupil total funding on indicators for HOLC districts for a subset of our analysis sample containing only D schools.[18] We first use the CBSA FE model described in our "Empirical Strategy" section. Our regressions show that D schools in HOLC B districts have, on average, $288 (SE = $584) more per-pupil total expenditures than districts mapped to HOLC D security ratings, while HOLC C districts have, on average, $1,332 (SE = $524) more than districts mapped to HOLC D security ratings. Differences between D vs. B are not statistically significant, but differences between D vs. C are statistically significant. Given the smaller HOLC D subsample, there is less HOLC A-D district variation at the CBSA level, making within-city comparisons few and far between. Running robustness checks and zooming out to the state and regional levels, our results are more aligned with the distributional plots in Figure 3. Namely, we can see that D schools in B HOLC districts have, on average, $1,755 (SE = $882) more per-pupil total expenditures than D HOLC districts, and D schools in C HOLC district have, on average, $1,369 (SE = $526) more per-pupil total expenditures than D HOLC districts. All differences are statistically significant.[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

[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. To facilitate a complete district-school comparison, we link our NERDS school and NCES F33 district level samples. 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 sample and NERDS school-level analysis sample, 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 region level but remain statistically significant for D vs. C differences.

to their most-in-need schools (i.e., D schools). 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 leads to 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.

6.1.4. 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 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). 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 a few robustness checks across samples. While the distribution of HOLC A-D grades across districts is slightly different between the 2017-18 cross-sectional and panel samples, we find no evidence to suggest that average outcomes by A-D HOLC grades vary by them (Figure S1.1, Table S1.1).[20] With these details in mind, let's discuss our results.

[Insert Figure 4 here]

First, to better understand how district funding changed over time by HOLC A-D grade, in Figure 4, we show A-D averages across time for each educational finance outcome, weighted

[20] 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 any of our F-33 educational finance variables.

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, C counterparts (Figure 5, Table 4). 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) relative to 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 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 Figure 5, 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.

[Insert Figure 5 here]

[Insert Table 4 here]

Finally, per-pupil federal revenues match much of the findings for per-pupil state revenues. However, there are some notable nuances. First, there is a large positive spike in per-pupil federal funding for the 2009-10 school year consistent with the surge in education funding from the American Recovery and Reinvestment Act of 2009 (ARRA), which allocated around $100 billion to the U.S. Department of Education who subsequently distributed it to states through the State Fiscal Stabilization Fund (SFSF), Title I – Part A, and the Individuals with Disabilities Education Act (IDEA) – Part B (U.S. Department of Education, 2009). In line with nationwide trends in per-pupil federal funding, in the decade that follows 2009-10 per-pupil federal funding decreases across all HOLC A-D grades although remains above historical averages from the late 1980s and 1990s. Next, while level differences in average per-pupil federal revenues existed before 2009-10, gaps between A districts and their lower B, C, and D counterparts clearly widen during this decade. Finally, unlike per-pupil state revenues, the D vs. A gap is not monotonically increasing. Although we see widening gaps from 1989-90 through 2009-10 between A vs. B, C, and D districts, they shrink in the years following the 2009-10 ARRA funding surge. This is exhibited for the D vs. A group in Figure 5.

6.2. School Racial Diversity

6.2.1. Current Outcomes

In Table 5, we see that both nationwide and by region, 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. Nationwide, for those schools mapped to HOLC D grades in our sample, 36 percent (SE = 0.011) of students are Black with differences greatest between A vs. D schools (-0.14, SE = 0.027) and smallest between C vs. D schools (-0.08, SE = 0.016). All these results are statistically significant. Regionally, we see similar patterns play out such that average shares of Black students in a school monotonically increases as one moves from the highest (i.e., A) to lowest (i.e., D) HOLC grade. Apart from the West region, all differences between D vs. A, B, and C schools are statistically significant. For the West region, only around 9 percent (SE = 0.012) of students in HOLC D schools are Black, and differences between D vs. A, B, and C schools are not statistically significant.

For percent non-White, we observe similar patterns to the above. Nationwide, for those schools mapped to D HOLC grades in our sample, 86 percent (SE =0.007) of students are non-White with differences greatest between A vs. D schools (-0.30, SE = 0.022) and smallest between C vs. D schools (-0.06, SE = 0.010). All results are statistically significant and are consistent across each U.S. Census Bureau region. By construction, results for percent White are inverse to percent non-White, such that nationwide for those schools mapped to D HOLC grades in our sample, 14 percent (SE = 0.007) of students are White with differences greatest between A vs. D schools (0.30, SE = 0.022) and smallest between C vs. D schools (0.06, SE = 0.010). Once again, results are all statistically significant and hold across all regions.

Our final school diversity outcome variable, and perhaps the most interesting when juxtaposed with the above, 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, SE = 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. Regionally, schools mapped to D HOLC grades are most diverse in the Northeast (0.50, SE = 0.002) and least diverse in the South (0.33, SE = 0.014). In general, the nationwide patterns described above hold across each region. However, all D vs. A, B, and C differences lack statistical significance for the Northeast region. In addition, D vs. C differences are not statistically significant for the South region.

[Insert Table 5 here]

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. Broadly, these takeaways are consistent with research on 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).

What might be driving these results? One explanation is that more money provides more opportunities to enact diversity mandates set by a district or city. While there are many examples of schools and districts announcing diversity mandates, their ability to reach diversity targets may depend on their available resources.[21] Student recruitment, student assignment plans, and targeted integration strategies all take time, staffing, and require financial commitments. As such, holding all else equal, one would predict that schools with less funding will have less diversity. However, our results indicate that HOLC D schools have the greatest shares of FRPL students, the most per-pupil expenditures, and the least diversity compared to HOLC A-C schools. These findings do not support the above narrative.

An alternative explanation, and one we think is more probable, is rooted in the understanding that wealthy families are less restricted spatially and less likely to be priced out of neighborhoods that act 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. What should one expect the status quo demographic composition to be, on average, in communities being served by low-performing schools? 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 &

[21] For example, see the New York State Board of Regents' recent announcement of diversity, equity, and inclusion initiatives for New York public schools (NYSED, 2021). In addition, many charters are adopting "diverse-by-design" approaches where diversity is explicitly stated as a core value or part of the school design (Potter & Quick, 2018).

Turner, 2008; Turner & Berube, 2009). Given this, we should expect low-performing schools to be less diverse, have higher concentrations of racial minorities, and have greater levels of low-income students. Consequently, our results are consistent with this theory and lend further evidence to the long and harmful half-life of historically discriminatory neighborhood policies on 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 nearly three decades and include 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 = 10,010) by over half (n_p = 4,693).[22] 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 a few 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.[23] While often small, these differences are almost always 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

[22] For the original 2018-19 cross-sectional data set we have n = 10,010. This translates to d = 2,037 and c = 141 distinct districts and CBSAs, respectively. For the 1988-89 through 2018-19 panel data set we have np = 4,693. This translates to d = 478 and c = 31 CBSAs, respectively.
[23] 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 each of our diversity outcomes. The exception being 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.

for the B security rating (Figure S2.1, Table S2.1). With these details in mind, we discuss the time series results below.

First, to better understand how school student demographics and racial diversity changed over time by HOLC A-D grade, in Figure 6 and Figure 7, 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 6. 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 percent Black from 1988-89 to 1998-99 reaching 24.8 and 26.5 percentage points, but this 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 flattens out over the final decade. The gap in average 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 by a few percentage points over the following decades to end at 7.4 and 6.5 percentage points, respectively.

For the percent White outcome, we see parallel downward sloping lines across time with negative gaps between D vs. A, B, and C counterparts remaining 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 6 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 that hover around 35 percentage points. For the percent Non-White outcome, we see the same patterns as the percent White outcome, except in reverse such that we observe upward sloping parallel lines across time with positive gaps between D vs. A, B, and C counterparts remaining relatively flat from 1988-89 through 2018-19.

[Insert Table 6 here]

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 nearly 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 relatively flat for those schools in the lowest-rated HOLC neighborhoods (i.e., C, D). Thus, by 2008-09 there are 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 once near-zero gaps in the diversity index in 1988-89 surpass -0.08 in 2018-19.

[Insert Figure 6 here]

[Insert Figure 7 here]

Our paper 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-redlined 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.

To complement the Simpson's Diversity Index results, we alternatively look at the Exposure Index for various race-group pairings. This is a well-known segregation index used frequently across the social sciences and differs from Simpson's Diversity Index by looking at exposure to different races via race-pair groupings. Looking at Figure S2.2 and Table S2.2, one can see an increase in White vs. Non-White exposure over time, driven largely by the increasing White-Hispanic Exposure Index, although the White-Asian Exposure Index does increase moderately over time, too. This further corroborates that U.S. schools, at least those in urban areas we consider in our sample, are beneficiaries of shifting U.S. demographics. Also, the decreasing White-Black Exposure Index closely mirrors nationwide trends and lends evidence to the White-Black resegregation narrative resulting from a reversal of desegregation efforts in the 1960s and 1970s.

Finally, over the three decades we consider in our time series analysis, we observe gaps between D vs. A, B, and C grades that are persistent and often growing 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.

6.3. School Student Performance

6.3.1. Current Outcomes

Nationwide and by region, there are virtually no statistically significant differences across HOLC A-D grades for the outcomes of average student learning and average student test score changes. When differences are statistically significant, they are small, with the largest statistically significant difference equating to just over 1/20th of a grade-level (Table 7). Each of their respective HOLC A-D grade averages, both nationwide and by region, is not statistically different from their respective grand means.

In contrast, the average student Math and ELA score outcome exhibit statistically significant differences between HOLC A-D grades both nationwide and across all regions. This is true for all HOLC D schools versus their higher A, B, 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.

[Insert Table 7 here]

These results tell us that while learning rates and changes in educational opportunity are, on average, the same across all HOLC A-D grades, overall educational opportunity is not. Specifically, those schools located in historically D assigned neighborhoods have less educational opportunity than those located in A, B, and C neighborhoods. For example, in Table 7 we see that nationwide A and D schools are separated by 0.64 SD units (SE = 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 (SE = 0.05) or about 2.25 grade levels in the West region and shrink to 0.54 SD units (SE = 0.10) or around 1.6 grade levels in the Northeast. Comparing B and D schools paints a similar picture as above, albeit somewhat muted, with schools separated by 0.34 SD units (SE = 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 (SE = 0.07) in the West and decreasing to as little as 0.30 SD units (SE = 0.06) in the South.

Finally, C and D schools show the greatest similarity of all D vs. A, B, and C comparisons with gaps shrinking to nearly single digits nationwide and across regions. Overall, C and D schools are separated by 0.13 SD units (SE = 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 (SE = 0.05), while the Northeast gap is smallest at 0.10 SD units (SE = 0.02). Both nationwide and by region, these differences in D vs. A, B and C educational opportunity are nearly all statistically significant.

7. DISCUSSION & 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 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, to the best of our knowledge, this effort has yet to extend to K-12 public school educational outcomes.

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, today, the least favorable overall and local district-finance outcomes relative to those mapped to higher-rated HOLC A, B, and C grades.[24] 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 past neighborhood inequality lingers well into the future. For example, those districts receiving redistributive funding to equalize local funding inequities are also those same districts serving 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 compared to 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 on average 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, much of these results

[24] Recall, 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.

persist across time, with overall positive time 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 appears to be 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. For the racial diversity outcomes, these differences are persistent and growing over time, 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 a more racially and ethnically diverse school-age population today. 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).[25] These countervailing forces could overwhelm, or at a minimum, limit some of the benefits that a more diverse U.S. population has on school racial diversity.

Finally, we also find that schools located today in historically-redlined have worse average ELA and math scores, but there is virtually no difference in both average learning rates and trends in test scores across A-D schools. Without a more exhaustive accounting of these

[25] 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.

trends over time, we cannot measure progress nor bring historical context to bear.[26] For example, if trends in educational opportunity favored D vs. A, B, and C schools over the past half century, 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 2009-18 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 SD 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.

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.

[26]While we have pooled 2009-2018 student performance SEDA data, we do not have access to longitudinal data on student performance outcomes for earlier decades.

Acknowledgments

We wish to thank Joshua Goodman, Martin West, Andrew Ho, Eric Taylor, 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

Dylan J. Lukes: First author; Conceptualization, Data Curation, Formal Analysis, Methodology, Project Administration – funding acquisition & grant application, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. Christopher H. Cleveland: Second author; Conceptualization, Writing – original draft, Writing – review & editing, Project Administration – funding identification.

Declaration of Competing Interests

This project received funding form the Bill & Melinda Gates Foundation. The research reported here was also supported in part by a training grant from the Institute of Education Sciences (R305B150010). Opinions expressed do not represent the views of the U.S. Department of Education. Finally, Christopher H. Cleveland is employed as a manager at the non-profit organization, Education Resource Strategies. None of the above-listed entities had any role in the study design, data collection, analysis and interpretation of data, in the writing of the report, and in the decision to submit the article for publication.

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].

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.

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. B., & Bartelt, D. W. (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.

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.

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.

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, Tomas, Brian Kisida, and Matthew M. Chingos. (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.

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.

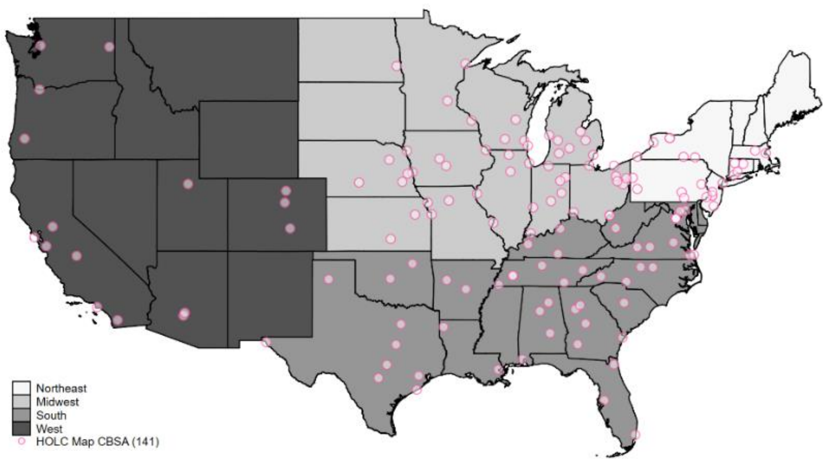
U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2020). Retrieved October 2, 2020, from https://nces.ed.gov/programs/edge/Home.

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.

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).

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	Time Series
Outcomes	Per-Pupil Revenue – Total, Local, State, Federal		% Black, White, Non-White, and Simpson's Diversity Index (1-D)		Average Test Scores, Student Learning, Trends in Test Scores	
Level	District		School		School	
Year	2017-2018	1989-2018	2018-2019	1989-2018	2009-2018	N/A
Schools	N/A		9,709	4,677	5,214	N/A
Districts	1,760	1,109	1,955	590	1,006	N/A
CBSAs	144	143	141	118	137	N/A
District Matches						
F33	1,760	1,109	1,283	581	891	N/A
Racial Diversity	1,283	581	1,955	590	1,001	N/A
SEDA	891	N/A	1,001	N/A	1,006	N/A

Notes: For a more detailed breakdown of outcome group samples by CBSA by HOLC A-D grade please refer to Appendix A where we provide school district counts and HOLC A-D percentages by CBSA across samples. Summary tables by outcome groups are also provided. These 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 Grades (USD 2018)

		U.S. Census Bureau Region				
		Nationwide	**Midwest**	**Northeast**	**South**	**West**
District PPR – Total, 2017-18 ($USD)	A	$1,546* [$822]	$1,891 [$1,306]	$2,343 [$1,490]	$1,973 [$2,484]	$29 [$647]
	B	$992* [$536]	$2,387* [$1,363]	$1,652 [$1,050]	$1,293 [$958]	-$1,082*** [$307]
	C	$1,504*** [$514]	$557 [$833]	$3,510* [$1,764]	$1,535* [$892]	$646** [$242]
	Constant	$14,402*** [$396]	$15,001*** [$491]	$19,863*** [$1,348]	$11,039*** [$812]	$14,629*** [$209]
	FE	Y	Y	Y	Y	Y
	R2	0.64	0.39	0.23	0.42	0.53
	F	4.08	3.72	2.79	1.13	29.51
	N	1,760	806	608	187	159
District PPR – Federal, 2017-18 ($USD)	A	-$722*** [$139]	-$687*** [$91]	-$941*** [$132]	-$812*** [$201]	-$255 [$386]
	B	-$391*** [$87]	-$522*** [$132]	-$373*** [$132]	-$349 [$313]	-$298** [$122]
	C	$253*** [$76]	$361*** [$101]	$238 [$223]	$249 [$190]	$182 [$151]
	Constant	$1,201*** [$62]	$1,182*** [$87]	$1,260*** [$150]	$1,201*** [$174]	$1,164*** [$127]
	FE	Y	Y	Y	Y	Y
	R2	0.32	0.24	0.33	0.47	0.37
	F	20.16	77.61	20.98	35.03	13.02
	N	1,760	806	608	187	159
District PPR – State, 2017-18 ($USD)	A	-$3,670*** [$752]	-$2,190*** [$582]	-$5,891*** [$1,903]	-$3,164*** [$1,002]	-$2,460* [$1,333]
	B	-$3,210*** [$776]	-$1,779*** [$380]	-$3,634* [$1,970]	-$3,149** [$1,213]	-$2,557*** [$693]
	C	-$461 [$495]	-$559 [$670]	$1,484 [$1,787]	-$1,114 [$804]	-$1,330 [$832]
	Constant	$7,960*** [$438]	$8,269*** [$510]	$10,149*** [$1,640]	$5,977*** [$743]	$9,342*** [$685]
	FE	Y	Y	Y	Y	Y
	R2	0.52	0.37	0.30	0.71	0.71
	F	8.75	9.28	31.95	7.68	4.59
	N	1,760	806	608	187	159
District PPR – Local, 2017-18 ($USD)	A	$5,937*** [$1,228]	$4,768*** [$1,340]	$9,175*** [$1,835]	$5,949** [$2,575]	$2,744** [$1,120]
	B	$4,592*** [$892]	$4,688*** [$1,594]	$5,659*** [$1,237]	$4,791** [$1,837]	$1,773* [$830]
	C	$1,713*** [$468]	$755** [$351]	$1,788 [$1,148]	$2,399 [$1,549]	$1,793** [$788]
	Constant	$5,240*** [$425]	$5,550*** [$343]	$8,454*** [$936]	$3,861*** [$1,421]	$4,123*** [$642]
	FE	Y	Y	Y	Y	Y
	R2	0.49	0.47	0.30	0.46	0.45
	F	11.56	9.97	8.51	3.82	3.35
	N	1,760	806	608	187	159

Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level
Note 2: 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 2017-18 total students. 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.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Table 3. School Per-Pupil Expenditures and 1935-1940 HOLC A-D Grades (USD 2018)

		U.S. Census Bureau Region				
		Nationwide	**Midwest**	**Northeast**	**South**	**West**
School PPE – Total, 2018-19 ($USD)	A	-$1,539* [$852]	$302 [$555]	-$3,022** [$1,416]	-$1,156*** [$214]	-$1,408** [$509]
	B	-$1,313** [$581]	$202 [$630]	-$2,309*** [$615]	-$606*** [$201]	-$1,558*** [$238]
	C	-$882*** [$334]	-$318 [$225]	-$1,579*** [$278]	-$295** [$143]	-$396 [$267]
	Constant	$16,890*** [$304]	$13,341*** [$200]	$21,871*** [$316]	$12,242*** [$91]	$14,285*** [$166]
	FE	Y	Y	Y	Y	Y
	R2	0.39	0.13	0.17	0.29	0.16
	F	3.67	1.01	141.54	10.55	39.04
	N	9,023	2,744	3,466	1,340	1,473
School PPE – Federal, 2018-19 ($USD)	A	-$602*** [$39]	-$535*** [$88]	-$653*** [$58]	-$636*** [$110]	-$490*** [$41]
	B	-$351*** [$29]	-$361*** [$56]	-$366*** [$47]	-$278*** [$73]	-$382*** [$77]
	C	-$141*** [$31]	-$155** [$67]	-$204*** [$35]	-$20 [$101]	-$62 [$57]
	Constant	$1,379*** [$18]	$1,402*** [$39]	$1,400*** [$26]	$1,479*** [$46]	$1,209*** [$32]
	FE	Y	Y	Y	Y	Y
	R2	0.14	0.17	0.14	0.17	0.05
	F	96.49	20.96	71.34	11.9	65.55
	N	8,573	2,363	3,466	1,339	1,405
School PPE – State & Local, 2018-19 ($USD)	A	-$1,131 [$860]	$572 [$663]	-$2,344 [$1,460]	-$643*** [$233]	-$971* [$468]
	B	-$1,010* [$590]	$599 [$751]	-$1,936*** [$622]	-$353** [$172]	-$1,193*** [$238]
	C	-$745** [$323]	-$126 [$221]	-$1,369*** [$292]	-$273** [$134]	-$372 [$218]
	Constant	$15,658*** [$297]	$12,110*** [$192]	$20,458*** [$326]	$10,652*** [$74]	$12,993*** [$145]
	FE	Y	Y	Y	Y	Y
	R2	0.41	0.14	0.18	0.3	0.17
	F	4.15	.52	109.79	4.3	29.43
	N	8,573	2,363	3,466	1,339	1,405

Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level
Note 2: 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 2018-19 total students. 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.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Figure 2. HOLC A-D School Distributions by HOLC A-D District Grades

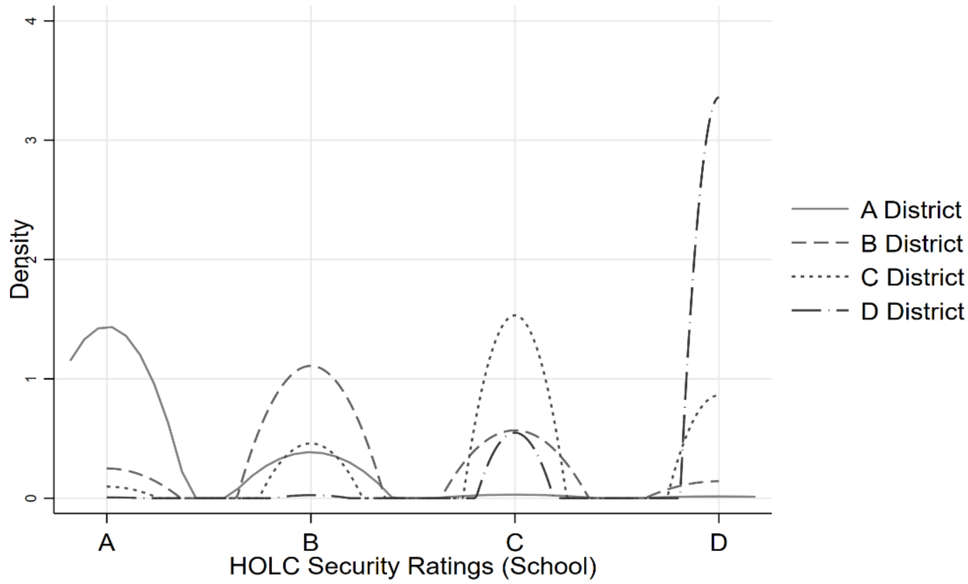


Figure 3. School Total PPE by HOLC A-D District Grades – Only "D" Schools

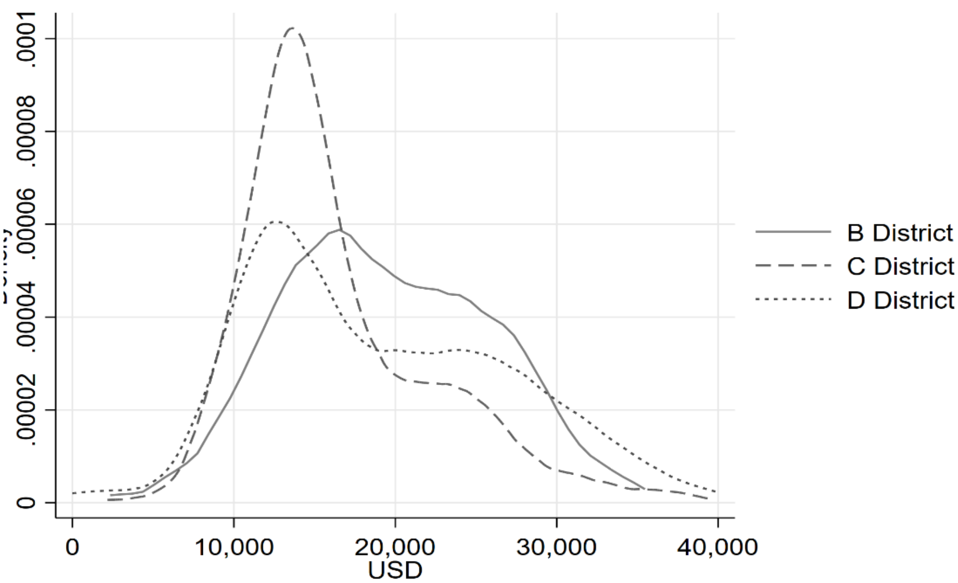
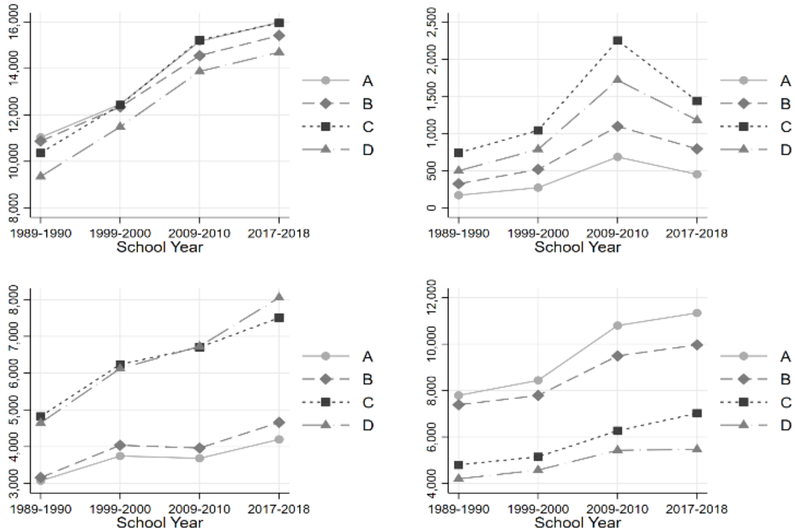


Table 4. Margins & Differences Over Time, Finance Outcomes, 1989-2018 (USD 2018)

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

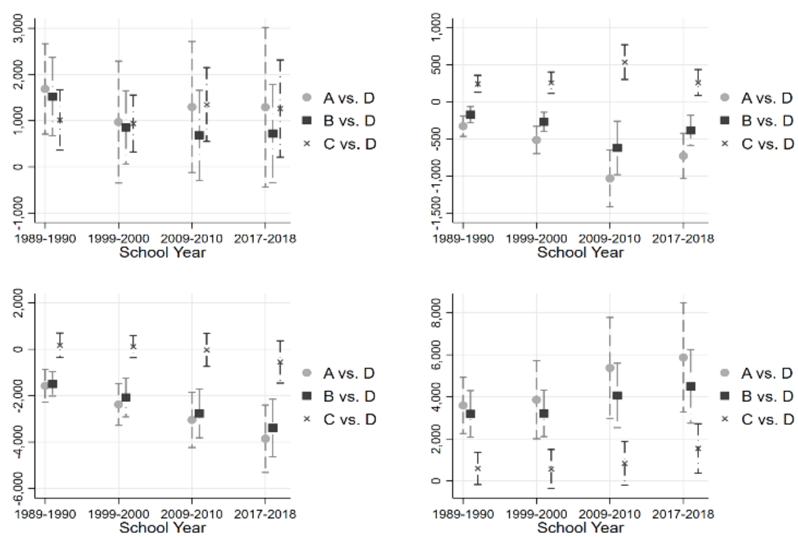
Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level
Note 2: For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. All regressions are weighted by 2017-18 total students. 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.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Figure 4. District Per-Pupil Revenue – Total, Overall & by Region, 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 5. HOLC Pairwise Comparisons Over Time, 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. Brackets represent 95% CIs. Standard errors are heteroskedasticity-consistent and clustered at the city-level. All dollars denominated in USD 2018.

Table 5. School Racial Diversity and 1935-1940 HOLC A-D Grades, 2018-2019

			U.S. Census Bureau Region			
		Nationwide	**Midwest**	**Northeast**	**South**	**West**
% Black	A	-0.14***	-0.23***	-0.11***	-0.17***	-0.01
		[0.027]	[0.037]	[0.03]	[0.045]	[0.03]
	B	-0.11***	-0.15***	-0.11***	-0.16***	-0.02
		[0.017]	[0.027]	[0.006]	[0.036]	[0.033]
	C	-0.08***	-0.12***	-0.08***	-0.1***	-0.01
		[0.016]	[0.022]	[0.017]	[0.034]	[0.012]
	Constant	0.36***	0.47***	0.35***	0.48***	0.09***
		[0.011]	[0.015]	[0.007]	[0.021]	[0.012]
	FE	Y	Y	Y	Y	Y
	R2	0.33	0.18	0.15	0.6	0.06
	F	13.74	18.36	137.6	7.36	.22
	N	9,709	3,031	3,595	1,453	1,630
% White	A	0.3***	0.31***	0.27***	0.29***	0.33***
		[0.022]	[0.034]	[0.032]	[0.059]	[0.024]
	B	0.18***	0.21***	0.18***	0.14***	0.21***
		[0.013]	[0.024]	[0.021]	[0.023]	[0.02]
	C	0.06***	0.09***	0.05***	0.05***	0.05***
		[0.01]	[0.025]	[0.014]	[0.017]	[0.013]
	Constant	0.14***	0.18***	0.14***	0.11***	0.11***
		[0.007]	[0.013]	[0.011]	[0.012]	[0.009]
	FE	Y	Y	Y	Y	Y
	R2	0.34	0.28	0.24	0.47	0.49
	F	86.13	40.14	53.58	14.31	68.32
	N	9,709	3,031	3,595	1,453	1,630
% Non-White	A	-0.3***	-0.31***	-0.27***	-0.29***	-0.33***
		[0.022]	[0.034]	[0.032]	[0.059]	[0.024]
	B	-0.18***	-0.21***	-0.18***	-0.14***	-0.21***
		[0.013]	[0.024]	[0.021]	[0.023]	[0.02]
	C	-0.06***	-0.09***	-0.05***	-0.05***	-0.05***
		[0.01]	[0.025]	[0.014]	[0.017]	[0.013]
	Constant	0.86***	0.82***	0.86***	0.89***	0.89***
		[0.007]	[0.013]	[0.011]	[0.012]	[0.009]
	FE	Y	Y	Y	Y	Y
	R2	0.34	0.28	0.24	0.47	0.49
	F	86.13	40.14	53.58	14.31	68.32
	N	9,709	3,031	3,595	1,453	1,630
Simpson's Diversity (1-D)	A	0.07***	0.06*	0.04	0.06*	0.2**
		[0.024]	[0.032]	[0.022]	[0.032]	[0.067]
	B	0.05***	0.05**	0.02	0.08***	0.1**
		[0.019]	[0.022]	[0.016]	[0.025]	[0.041]
	C	0.03***	0.04***	0.01	0.02	0.08***
		[0.011]	[0.014]	[0.008]	[0.025]	[0.011]
	Constant	0.42***	0.39***	0.5***	0.33***	0.35***
		[0.009]	[0.007]	[0.002]	[0.014]	[0.014]
	FE	Y	Y	Y	Y	Y
	R2	0.29	0.29	0.07	0.34	0.33
	F	3.8	9.36	8.48	5.99	32.59
	N	9,709	3,031	3,595	1,453	1,630

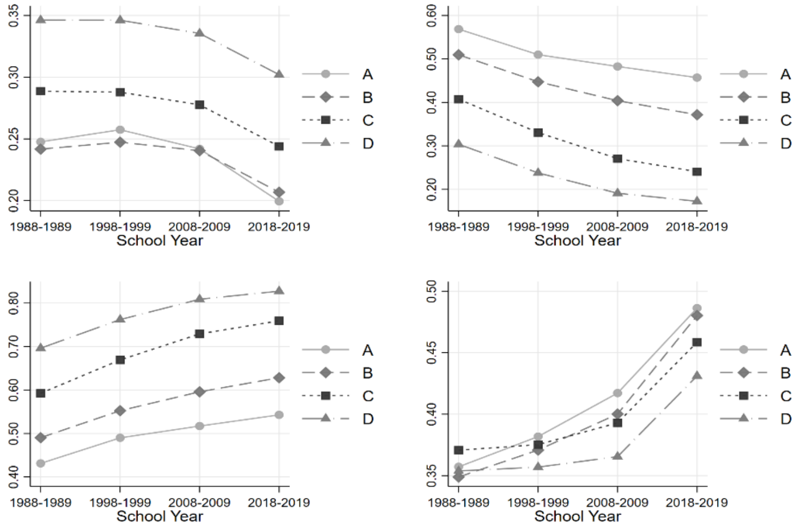
Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level
Note 2: For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. All regressions are weighted by 2018-19 total students. All models use city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Table 6. 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 [0.027]	0.24 [0.009]	0.29 [0.008]	0.35 [0.021]	-0.1** [0.042]	-0.1*** [0.029]	-0.06** [0.029]
	1998-99	0.26 [0.029]	0.25 [0.01]	0.29 [0.007]	0.35 [0.021]	-0.09** [0.043]	-0.1*** [0.029]	-0.06** [0.026]
	2008-09	0.24 [0.026]	0.24 [0.009]	0.28 [0.008]	0.34 [0.02]	-0.09** [0.041]	-0.1*** [0.027]	-0.06** [0.027]
	2018-19	0.20 [0.022]	0.21 [0.007]	0.24 [0.007]	0.30 [0.018]	-0.1*** [0.034]	-0.1*** [0.023]	-0.06** [0.024]
% White	1988-89	0.57 [0.027]	0.51 [0.015]	0.41 [0.005]	0.3 [0.011]	0.26*** [0.028]	0.21*** [0.025]	0.1*** [0.014]
	1998-99	0.51 [0.028]	0.45 [0.016]	0.33 [0.007]	0.24 [0.011]	0.27*** [0.028]	0.21*** [0.024]	0.09*** [0.014]
	2008-09	0.48 [0.027]	0.4 [0.016]	0.27 [0.007]	0.19 [0.011]	0.29*** [0.029]	0.21*** [0.023]	0.08*** [0.014]
	2018-19	0.46 [0.024]	0.37 [0.013]	0.24 [0.006]	0.17 [0.01]	0.28*** [0.028]	0.2*** [0.02]	0.07*** [0.013]
% Non-White	1988-89	0.43 [0.027]	0.49 [0.015]	0.59 [0.005]	0.7 [0.011]	-0.26*** [0.028]	-0.21*** [0.025]	-0.1*** [0.014]
	1998-99	0.49 [0.028]	0.55 [0.016]	0.67 [0.007]	0.76 [0.011]	-0.27*** [0.028]	-0.21*** [0.024]	-0.09*** [0.014]
	2008-09	0.52 [0.027]	0.6 [0.016]	0.73 [0.007]	0.81 [0.011]	-0.29*** [0.029]	-0.21*** [0.023]	-0.08*** [0.014]
	2018-19	0.54 [0.024]	0.63 [0.013]	0.76 [0.006]	0.83 [0.01]	-0.28*** [0.028]	-0.2*** [0.02]	-0.07*** [0.013]
Simpson's Diversity (1-D)	1988-89	0.36 [0.019]	0.35 [0.006]	0.37 [0.009]	0.35 [0.017]	<0.01 [0.03]	<0.01 [0.02]	0.02 [0.026]
	1998-99	0.38 [0.02]	0.37 [0.008]	0.38 [0.007]	0.36 [0.017]	0.02 [0.033]	0.01 [0.022]	0.02 [0.023]
	2008-09	0.42 [0.023]	0.4 [0.01]	0.39 [0.005]	0.37 [0.015]	0.05 [0.036]	0.03 [0.023]	0.03 [0.017]
	2018-19	0.49 [0.021]	0.48 [0.01]	0.46 [0.006]	0.43 [0.011]	0.06* [0.03]	0.05*** [0.018]	0.03** [0.013]

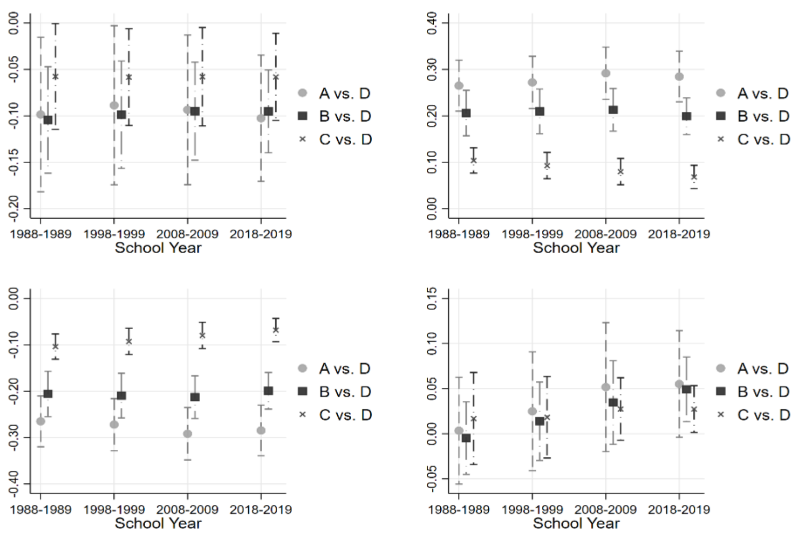
Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level
Note 2: For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. All regressions are weighted by 2018-19 total students. 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.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Figure 6. 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.

Figure 7. HOLC Pairwise Comparisons Over Time, 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. Brackets represent 95% confidence intervals. Standard errors are heteroskedasticity-consistent and clustered at the city-level.

Table 7. School Student Performance and 1935-1940 HOLC A-D Grades, Pooled 2008-2019

		U.S. Census Bureau Region				
		Nationwide	**Midwest**	**Northeast**	**South**	**West**
Avg. Student Math & ELA Scores	A	0.643***	0.652***	0.635***	0.541***	0.753***
		[0.04]	[0.06]	[0.064]	[0.102]	[0.054]
	B	0.342***	0.306***	0.336***	0.296***	0.46***
		[0.028]	[0.055]	[0.044]	[0.064]	[0.072]
	C	0.127***	0.142***	0.103***	0.153***	0.126***
		[0.015]	[0.035]	[0.024]	[0.045]	[0.03]
	Constant	-0.365***	-0.459***	-0.225***	-0.366***	-0.461***
		[0.013]	[0.023]	[0.021]	[0.034]	[0.024]
	FE	Y	Y	Y	Y	Y
	R2	0.27	0.17	0.28	0.24	0.27
	F	91.85	40.53	35.15	10.66	79.36
	N	6,890	2,181	2,576	1,042	1,091
Avg. Student Learning (Annual)	A	0.006	0.011	0.023**	-0.011	-0.009
		[0.006]	[0.008]	[0.009]	[0.012]	[0.009]
	B	0.008*	0.003	0.013	0.003	0.012**
		[0.004]	[0.007]	[0.008]	[0.007]	[0.006]
	C	0.005*	0.007*	0.005	0.006	-0.002
		[0.003]	[0.004]	[0.006]	[0.007]	[0.003]
	Constant	0.007***	0.016***	-0.004	-0.005	0.02***
		[0.002]	[0.002]	[0.005]	[0.004]	[0.002]
	FE	Y	Y	Y	Y	Y
	R2	0.14	0.14	0.09	0.24	0.05
	F	1.51	2.11	6.9	1.71	7.97
	N	5,193	1,854	1,617	875	847
Avg. Student Test Score Change	A	0.01***	0.014***	0.014***	0.006	0.003
		[0.003]	[0.005]	[0.004]	[0.006]	[0.005]
	B	0.007***	0.006	0.004	0.01**	0.008
		[0.002]	[0.004]	[0.005]	[0.005]	[0.006]
	C	0.004***	0.007***	0.001	0.006*	0
		[0.001]	[0.002]	[0.002]	[0.003]	[0.002]
	Constant	-0.003***	-0.01***	-0.002	-0.01***	0.012***
		[0.001]	[0.002]	[0.002]	[0.002]	[0.002]
	FE	Y	Y	Y	Y	Y
	R2	0.19	0.15	0.18	0.27	0.08
	F	5.1	5.58	11.11	1.53	2.85
	N	6,183	2,032	2,135	976	1,040

Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level.
Note 2: 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.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

APPENDIX A.

Table A1. Finance Sample Overview: All 2017-18 U.S. Public School Districts vs. HOLC Analysis 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%

Note: First line of each variable includes all active 2017-18 U.S. public primary and secondary districts. 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 districts for the HOLC A-D analysis sample. Districts were only included in this sample if they were matched to 1935-1940 HOLC A-D maps. Shares represent match rates for "Districts" and "CBSAs" variables and sample representativeness for all others.

Table A2. Diversity Sample Overview: All 2018-19 U.S. Public Schools vs. HOLC Analysis 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%

Note: First line of each variable includes all active 2018-19 U.S. public primary and secondary schools. Only those schools with non-missing and non-zero student enrollments included. The second line represents all 2018-19 U.S. public primary and secondary schools for the HOLC A-D "Student Diversity" analysis sample. Schools were only included in this sample if they were matched to 1935-1940 HOLC A-D maps. Shares represent match rates with HOLC A-D maps for "Schools", "Districts" and "CBSAs" variables and simply sample representativeness for all others variables.

Table A3. SEDA Sample Overview: All 2017-18 U.S. Public School Districts vs. HOLC Analysis 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%

Note: 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 A4. CBSA Summary Statistics by 1935-1940 HOLC Grade & Outcome Group

	Outcome Categories														
	F33 District Finance (A-D%)					**School Diversity (A-D%)**					**School Performance (A-D%)**				
	A%	*B%*	*C%*	*D%*	*N*	*A%*	*B%*	*C%*	*D%*	*N*	*A%*	*B%*	*C%*	*D%*	*N*
Akron, OH	4	20	72	4	25	13	32	49	6	17	13	30	50	7	14
Albany-Schenectady-Troy, NY	4	7	74	15	27	12	22	48	18	18	14	21	47	19	15
Allentown-Bethlehem-Easton, PA-NJ	0	25	75	0	4	0	80	20	0	1	0	80	20	0	1
Altoona, PA	0	25	75	0	4	0	14	71	14	2	0	20	60	20	1
Amarillo, TX	0	0	100	0	1	22	9	22	48	1	25	10	25	40	1
Asheville, NC	0	67	33	0	3	0	33	58	8	3	0	25	63	13	3
Atlanta-Sandy Springs-Roswell, GA	0	0	83	17	6	0	18	37	45	6	0	21	35	44	4
Atlantic City-Hammonton, NJ	0	27	64	9	11	0	12	65	24	5	0	15	62	23	4
Augusta-Richmond County, GA-SC	0	0	100	0	2	0	0	20	80	1	0	0	20	80	1
Austin-Round Rock, TX	17	33	0	50	6	13	29	13	45	6	17	30	9	43	4
Baltimore-Columbia-Towson, MD	0	33	67	0	3	6	32	32	30	3	6	32	32	29	3
Battle Creek, MI	0	0	50	50	6	0	0	67	33	6	0	0	86	14	4
Bay City, MI	0	0	40	60	5	0	10	20	70	3	0	13	25	63	3
Beaumont-Port Arthur, TX	0	60	40	0	5	16	58	16	11	4	7	64	14	14	4
Binghamton, NY	0	43	57	0	7	7	67	27	0	4	8	69	23	0	4
Birmingham-Hoover, AL	11	0	33	56	9	4	9	35	52	8	6	12	42	39	6
Boston-Cambridge-Newton, MA-NH	1	13	64	21	67	3	13	55	29	55	4	15	54	28	50
Bridgeport-Stamford-Norwalk, CT	38	38	13	13	8	24	41	18	18	5	15	46	23	15	3
Buffalo-Cheekt.-Niag. Falls, NY	10	45	40	5	20	5	33	54	8	17	8	33	51	8	11
Canton-Massillon, OH	0	30	60	10	10	5	36	41	18	7	6	41	41	12	5
Charleston, WV	0	0	100	0	1	13	25	50	13	1	14	29	57	0	1
Charlotte-Concord-Gastonia, NC-SC	0	0	50	50	2	13	7	27	53	2	14	7	29	50	2
Chattanooga, TN-GA	0	0	50	50	2	12	12	53	24	1	13	13	53	20	1
Chicago-Naperville-Elgin, IL-IN-WI	2	22	50	26	140	1	11	50	38	84	2	11	51	37	72
Cincinnati, OH-KY-IN	18	18	65	0	17	23	26	43	9	12	15	27	46	12	11
Cleveland-Elyria, OH	2	16	55	27	93	5	17	50	28	81	5	18	51	26	67
Columbia, SC	0	0	100	0	1	0	29	29	43	3	0	20	40	40	2
Columbus, GA-AL	0	0	100	0	1	0	0	67	33	1	0	0	50	50	1
Columbus, OH	11	21	39	29	28	12	25	42	21	27	12	26	48	14	18
Concord, NH	0	0	100	0	1	0	0	0	0	0	0	0	0	0	0
Dallas-Fort Worth-Arlington, TX	8	0	67	25	12	7	19	51	23	31	9	25	45	22	12
Davenport-Moline-Rock Island, IA-IL	0	50	50	0	2	0	42	47	11	1	0	46	46	8	1
Dayton, OH	9	5	41	45	22	17	9	37	37	17	14	11	39	36	13
Decatur, IL	0	0	50	50	2	10	0	60	30	3	25	0	50	25	1
Denver-Aurora-Lakewood, CO	0	0	89	11	9	4	9	43	44	4	4	13	46	38	4
Des Moines-West Des Moines, IA	0	0	100	0	5	25	20	35	20	1	21	21	38	21	1
Detroit-Warren-Dearborn, MI	1	6	53	40	116	5	10	50	35	105	5	14	51	31	94
Dubuque, IA	0	0	100	0	1	22	56	11	11	1	25	50	13	13	1
Duluth, MN-WI	0	0	50	50	2	44	11	22	22	2	38	13	25	25	2
Durham-Chapel Hill, NC	0	0	75	25	8	7	7	67	20	7	9	9	64	18	6
El Paso, TX	0	0	50	50	2	8	25	33	33	2	13	25	38	25	2
Elmira, NY	0	0	100	0	3	11	33	33	22	3	0	33	50	17	1
Erie, PA	0	57	43	0	7	0	71	18	12	4	0	67	20	13	4
Evansville, IN-KY	0	0	67	33	3	7	0	67	27	2	8	0	62	31	2
Flint, MI	0	10	10	80	10	0	33	11	56	6	0	50	0	50	3
Fort Wayne, IN	0	0	100	0	3	0	10	60	30	2	0	11	67	22	2
Fresno, CA	0	0	50	50	2	0	4	68	28	4	0	5	70	25	1
Grand Rapids-Wyoming, MI	0	15	80	5	20	3	19	73	6	18	2	20	73	6	15
Greensboro-High Point, NC	0	0	100	0	1	0	29	43	29	1	0	50	25	25	1
Harrisburg-Carlisle, PA	11	11	44	33	9	0	36	14	50	6	0	42	17	42	5
Hartford-West Hartford-East Hartford, CT	7	57	29	7	14	10	29	47	14	9	10	27	48	15	7
Houston-Woodlands-Sugar Land, TX	0	33	33	33	3	10	31	42	17	11	6	33	44	17	8
Huntington-Ashland, WV-KY-OH	0	0	100	0	2	17	25	58	0	3	18	27	55	0	2
Indianapolis-Carmel-Anderson, IN	0	2	65	33	46	2	10	61	28	49	2	10	58	29	29
Jackson, MI	0	0	100	0	7	16	16	63	5	6	20	10	60	10	3
Jackson, MS	0	0	67	33	3	8	23	31	38	2	0	30	30	40	2
Jacksonville, FL	0	0	100	0	1	0	29	4	68	1	0	26	4	70	1
Johnstown, PA	0	20	60	20	5	0	33	67	0	2	0	33	67	0	2
Kalamazoo-Portage, MI	17	0	83	0	6	0	6	83	11	4	0	11	67	22	2
Kansas City, MO-KS	0	13	33	54	24	1	5	32	62	27	1	5	32	62	20
Knoxville, TN	0	0	100	0	1	4	9	48	39	1	7	7	33	53	1
Lancaster, PA	50	50	0	0	2	0	100	0	0	1	0	100	0	0	1
Lansing-East Lansing, MI	0	36	64	0	11	3	15	79	3	10	4	14	82	0	9
Lexington-Fayette, KY	0	0	100	0	1	17	33	42	8	1	22	22	44	11	1
Lima, OH	0	0	100	0	5	0	33	44	22	1	0	43	57	0	1
Lincoln, NE	0	100	0	0	3	27	29	44	0	3	27	31	42	0	1
Little Rock-North Little Rock-Conway, AR	20	0	40	40	5	24	14	14	48	6	14	14	21	50	4
Los Angeles-Long Beach-Anah., CA	3	17	68	12	65	4	11	51	35	227	5	13	50	32	47
Louisville/Jefferson County, KY-IN	0	0	100	0	2	4	13	37	46	2	3	18	42	37	1
Lynchburg, VA	0	0	100	0	2	0	0	67	33	2	0	0	100	0	1
Macon, GA	0	0	0	0	0	0	13	38	50	3	0	0	40	60	2
Macon-Bibb County, GA	0	0	50	50	2	0	0	0	0	0	0	0	0	0	0
Madison, WI	0	0	60	40	5	17	11	67	6	2	27	18	55	0	1
Manchester-Nashua, NH	0	0	67	33	6	7	7	73	13	3	9	9	82	0	2
Memphis, TN-MS-AR	0	0	100	0	1	0	19	35	46	2	0	14	40	46	2
Miami-Fort Lauderdale-West Palm Beach, FL	0	0	100	0	1	6	7	32	54	1	7	9	32	52	1
Milwaukee-Waukesha-West Allis, WI	6	12	48	33	33	4	13	52	31	23	4	12	54	30	17
Minneapolis-St. Paul-Blmt., MN-WI	3	34	29	33	58	6	33	36	26	55	7	36	35	22	43
Mobile, AL	0	0	100	0	1	0	0	29	71	1	0	0	29	71	1
Monroe, MI	0	0	100	0	3	0	0	0	0	0	0	0	0	0	0
Montgomery, AL	0	0	25	75	4	0	12	35	53	5	0	10	40	50	1

Muncie, IN	0	50	50	0	2	0	100	0	0	1	0	0	0	0	0
Muskegon, MI	0	25	38	38	8	0	40	20	40	5	0	50	0	50	4
Nashville-Davidson-Murf.-Frnk., TN	0	0	33	67	3	6	14	17	63	2	7	14	19	60	1
New Castle, PA	20	20	40	20	5	0	0	100	0	1	0	0	100	0	1
New Haven-Milford, CT	0	14	43	43	14	4	12	63	21	8	4	14	63	20	8
New Orleans-Metairie, LA	0	7	39	54	41	0	7	52	42	44	0	7	48	45	33
NYC-Newark-Jersey, NY-NJ-PA	2	19	39	40	424	3	16	40	41	396	3	17	42	38	275
Ogden-Clearfield, UT	0	0	25	75	4	10	19	43	29	4	7	20	33	40	3
Oklahoma City, OK	30	0	50	20	10	19	22	46	14	7	14	21	48	17	5
Omaha-Council Bluffs, NE-IA	11	67	22	0	9	13	53	19	15	7	14	52	21	14	5
Oshkosh-Neenah, WI	0	0	100	0	1	0	0	73	27	1	0	0	73	27	1
Peoria, IL	0	0	63	38	8	0	7	60	33	3	0	9	55	36	3
Phil.-Camden-Wilm., PA-NJ-DE-MD	6	27	39	29	108	4	33	25	38	98	4	33	26	36	85
Phoenix-Mesa-Scottsdale, AZ	10	23	42	26	31	7	35	33	26	27	7	31	31	31	18
Pittsburgh, PA	5	18	47	29	38	6	36	33	26	23	5	34	34	26	22
Platteville, WI	0	0	0	100	1	0	0	0	0	0	0	0	0	0	0
Portland-Vanc.-Hillsboro, OR-WA	33	17	50	0	6	3	22	61	14	4	3	22	63	12	4
Portsmouth, OH	0	0	33	67	6	0	0	100	0	2	0	0	100	0	2
Providence-Warwick, RI-MA	4	33	46	17	24	5	29	57	9	20	6	31	56	8	12
Pueblo, CO	0	50	50	0	2	18	41	24	18	2	15	38	31	15	1
Racine, WI	0	0	100	0	1	0	31	46	23	1	0	40	40	20	1
Richmond, VA	0	0	100	0	2	4	21	29	46	1	5	10	40	45	1
Roanoke, VA	0	0	25	75	4	0	11	44	44	3	0	14	43	43	2
Rochester, MN	0	100	0	0	1	38	13	50	0	1	33	33	33	0	1
Rochester, NY	5	50	40	5	20	4	23	49	23	17	5	22	47	25	13
Rockford, IL	0	0	67	33	3	0	6	39	56	2	0	7	36	57	2
Sacramento-Rsvll.--Arden-Arc., CA	0	0	67	33	3	9	9	64	18	2	11	11	67	11	2
Saginaw, MI	0	17	50	33	6	0	15	80	5	7	0	20	80	0	3
Salt Lake City, UT	0	11	33	56	9	6	27	24	42	7	6	26	26	42	7
San Antonio-New Braunfels, TX	13	20	47	20	15	8	26	31	35	25	9	20	38	33	11
San Diego-Carlsbad, CA	0	14	43	43	7	4	11	34	52	24	4	14	30	52	2
San Francisco-Oakland-Hayward, CA	22	11	50	17	18	5	18	39	38	54	6	21	41	32	8
San Jose-Sunnyvale-Santa Clara, CA	0	0	50	50	10	5	11	42	42	9	9	9	45	36	3
Savannah, GA	0	0	100	0	1	13	13	50	25	1	0	20	40	40	1
Scranton--Wilkes-Barre-Hazl., PA	14	57	0	29	7	86	14	0	0	2	100	0	0	0	2
Seattle-Tacoma-Bellevue, WA	0	0	83	17	12	1	32	41	26	9	1	35	49	15	5
Shreveport-Bossier City, LA	0	0	100	0	2	18	9	36	36	2	18	9	36	36	2
Sioux City, IA-NE-SD	0	14	14	71	7	5	10	19	67	1	6	13	25	56	1
South Bend-Mishawaka, IN-MI	0	0	83	17	6	10	10	52	28	5	9	9	59	23	4
Spokane-Spokane Valley, WA	0	0	83	17	6	2	20	56	22	3	4	30	59	7	2
Springfield, IL	0	50	50	0	2	6	25	13	56	1	8	15	8	69	1
Springfield, MA	0	14	86	0	7	5	14	67	14	4	7	7	64	21	4
Springfield, MO	0	0	50	50	2	0	14	48	38	3	0	20	47	33	1
Springfield, OH	0	17	67	17	6	0	30	50	20	4	0	25	50	25	3
St. Joseph, MO-KS	0	0	0	100	1	0	0	0	100	1	0	0	0	100	1
St. Louis, MO-IL	9	25	36	30	44	11	27	29	33	35	12	31	28	29	30
Stockton-Lodi, CA	0	0	100	0	1	0	14	43	43	1	0	17	33	50	1
Syracuse, NY	8	25	50	17	12	14	34	29	23	11	15	33	30	22	7
Tampa-St. Petersburg-Clearwater, FL	0	0	100	0	2	2	10	28	61	2	2	11	24	63	2
Terre Haute, IN	0	0	100	0	2	14	0	43	43	2	17	0	33	50	1
Toledo, OH	4	30	61	4	23	4	25	62	9	19	4	27	59	10	18
Topeka, KS	0	0	33	67	3	8	23	38	31	2	9	27	36	27	1
Torrington, CT	0	0	100	0	1	0	0	100	0	1	0	0	100	0	1
Trenton, NJ	14	14	57	14	7	7	21	55	17	8	5	23	55	18	7
Tulsa, OK	50	0	50	0	2	29	12	18	41	2	31	13	19	38	2
Utica-Rome, NY	0	33	33	33	6	13	13	60	13	3	17	8	67	8	2
Virginia Beach-Norfolk-Newport News, VA-NC	0	0	71	29	7	3	14	31	53	6	4	20	36	40	5
Waco, TX	0	25	50	25	4	0	47	20	33	5	0	45	27	27	4
Waterloo-Cedar Falls, IA	0	0	100	0	1	10	40	20	30	1	13	38	25	25	1
Wheeling, WV-OH	0	0	50	50	2	14	14	14	57	2	14	14	14	57	2
Wichita, KS	0	0	100	0	1	12	12	4	72	1	14	14	5	67	1
Winston-Salem, NC	0	0	100	0	2	7	7	50	36	2	0	8	54	38	2
York-Hanover, PA	0	25	63	13	8	0	13	63	25	4	0	17	83	0	3
Youngstown-Warren-Board., OH-PA	0	14	54	32	28	0	26	42	32	19	0	21	42	38	15
Total Districts			2,135					2,037					1,371		

Figure A1. 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 $3600-$6000 & 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	$	% 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	% change	$	% change
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. $5500-$7500 & 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 (3 yr.) a. HOLC 2 b. Institutions Few
6. MORTGAGE FUNDS: Ample 7. TOTAL TAX RATE PER $1000 (1937-1938) $54.95
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	$ 500-1000	% change	$	% 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	$	% 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			

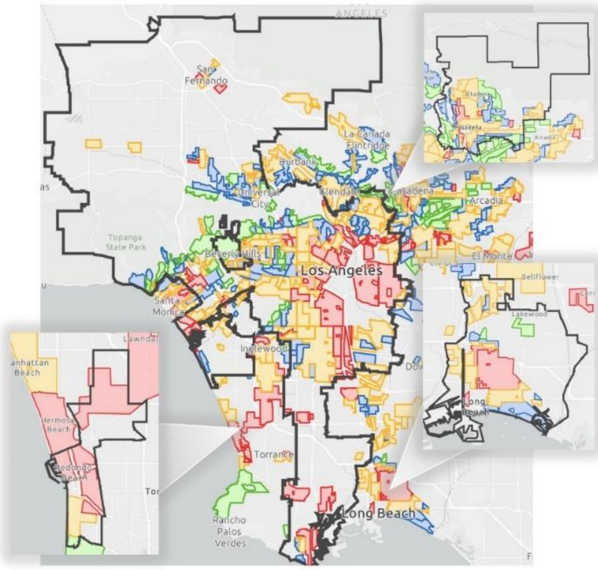
3. NEW CONSTRUCTION (past yr.) No. 0 Type & Price -- How Selling --
4. OVERHANG OF HOME PROPERTIES: a. HOLC 2 b. Institutions Few
5. SALE OF HOME PROPERTIES (3 yr.) a. HOLC 2 b. Institutions Few
6. MORTGAGE FUNDS: None 7. TOTAL TAX RATE PER $1000 (1937-1938) $52.70
8. DESCRIPTION AND CHARACTERISTICS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

Terrain: Level. No construction or flood hazards. Land improved 60%. 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 3/25/39

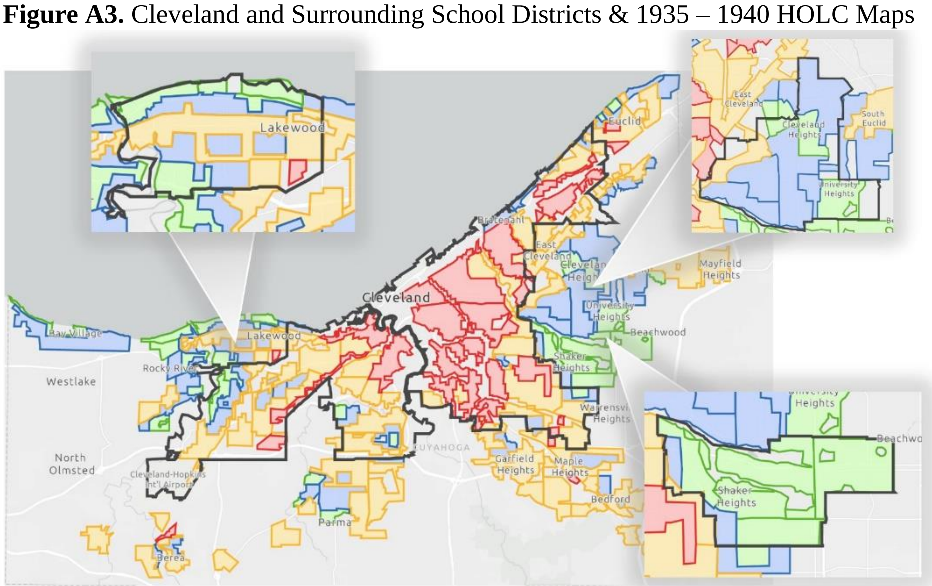
349

Figure A2. 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 A3. Cleveland and Surrounding School Districts & 1935 – 1940 HOLC Maps

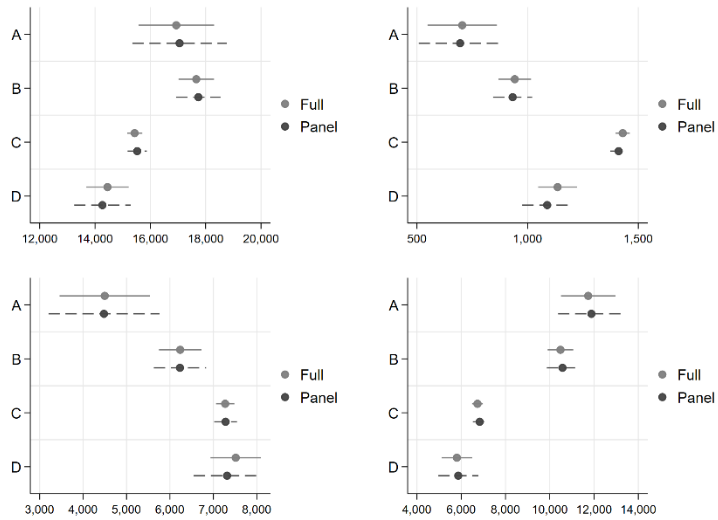


Notes: "Best" (A, outlined in green), "Still Desirable" (B, outlined in blue), "Definitely Declining" (C, outlined in yellow), to "Hazardous" (D, outlined in red) [center] Cleveland Municipal; [top left] Lakewood City; [top right] Cleveland Heights-University Heights; [bottom right] Shaker Heights

ONLINE APPENDIX: SUPPLEMENTAL TABLES & FIGURES

Appendix S1. District Finance Outcomes

Figure S1.1. HOLC A-D Coefficient Plots, Full vs. Panel Sample, 2017-18, Finance Outcomes (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.

Table S1.1. HOLC A-D Means, SEs and Differences, Full vs. Panel Sample, Finance Outcomes (USD 2018)

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

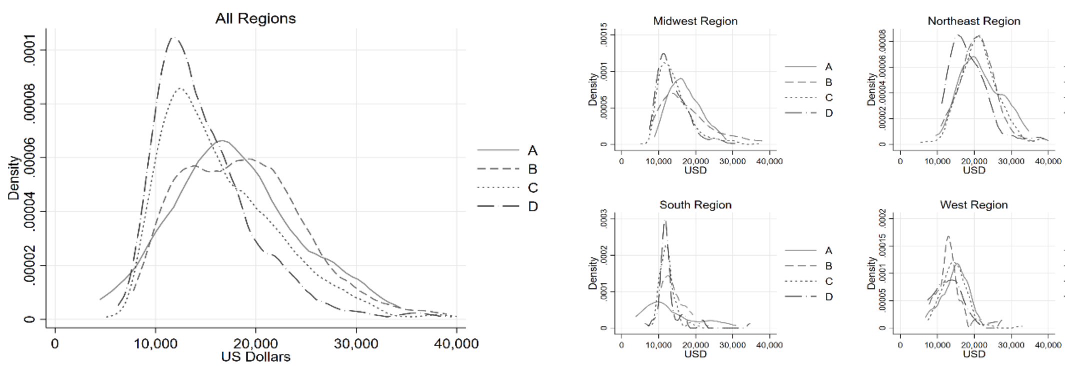
Notes: Means are HOLC A-D weighted averages from regressions of respective finance outcomes on HOLC B-D indicators. Standard errors are Huber-White heteroskedasticity-consistent and are in brackets. Regressions are run using both the 2017-18 full sample and the 2017-18 data from the panel sample. The 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 dp = 1,109 districts.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Table S1.2. Kolmogorov-Smirnov Equality of Distribution Test, Finance Outcomes

		U.S. Census Bureau Region[b]				
		Overall	**Midwest**	**Northeast**	**South**	**West**
District PPR – Total, 2017-18 ($USD)	D vs. A[a]	-0.28***	-0.44***	-0.29**	-0.17	-0.27
	D vs. B	-0.33***	-0.28***	-0.30***	-0.36**	-0.17
	D vs. C	-0.14***	-0.10*	-0.31***	-0.15	-0.26**
District PPR – Federal, 2017-18 ($USD)	A vs. D	0.51***	0.48***	0.59***	0.54***	0.61***
	B vs. D	0.38***	0.38***	0.43***	0.19	0.29
	C vs. D	0.11***	0.11**	0.22***	0.19*	0.12
District PPR – State, 2017-18 ($USD)	A vs. D	0.39***	0.55***	0.22*	0.23	0.30
	B vs. D	0.18***	0.25***	0.04	0.24	0.18
	C vs. D	0.05	0.16***	<0.01	0.13	0.07
District PPR – Local, 2017-18 ($USD)	D vs. A	-0.41***	-0.60***	-0.36***	-0.31	-0.42*
	D vs. B	-0.38***	-0.40***	-0.20***	-0.4**	-0.26
	D vs. C	-0.16***	-0.22***	-0.12	-0.16	-0.22

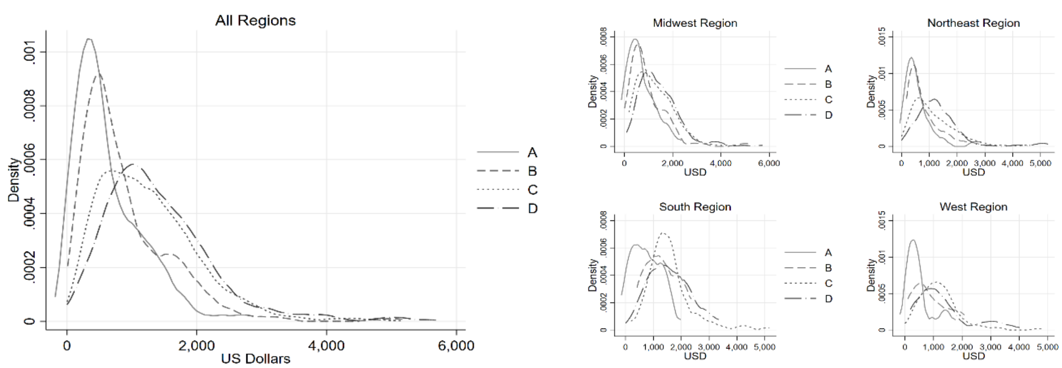
Notes: a – Tests the null hypothesis that the first HOLC grade distribution contains smaller values than the second HOLC grade distribution. For example, D vs. A test the null hypothesis that D distribution contains smaller values than A distribution; b – The values seen here, both for overall and each region, represent the largest difference between distributions in the direction as specified from the "Group" row input.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Figure S1.2. District Per-Pupil Revenue – Total, Overall & by Region, 2017-18



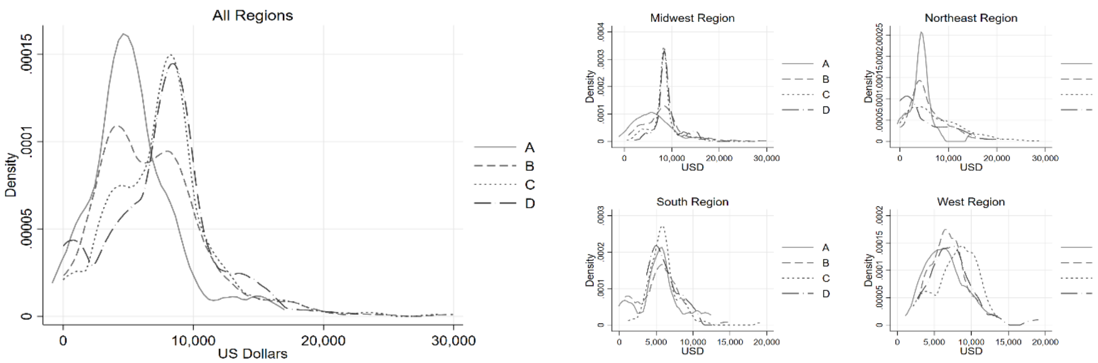
Notes: Epanechnikov kernel used for univariate kernel density estimation. Per-pupil total revenues are restricted to <$40K to remove outliers.

Figure S1.3. District Per-Pupil Revenue – Federal, Overall & by Region, 2017-18



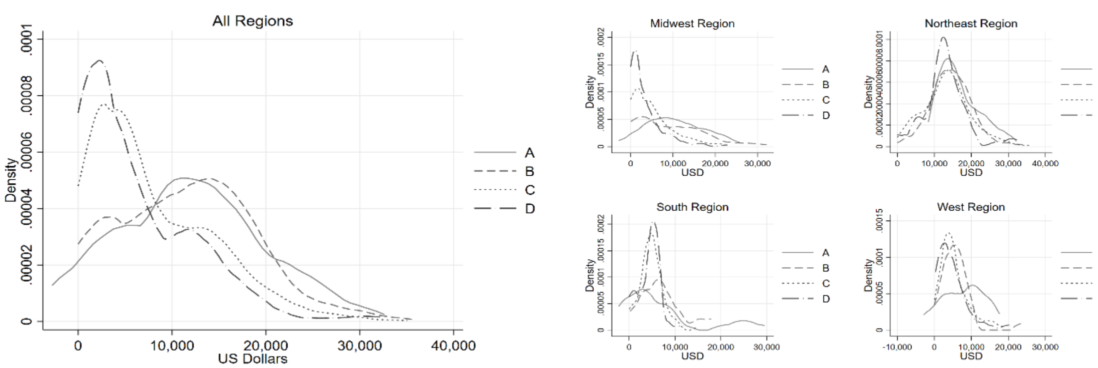
Notes: Epanechnikov kernel used for univariate kernel density estimation. Per-pupil federal revenues are restricted to <$6K to remove outliers.

Figure S1.4. District Per-Pupil Revenue – State, Overall & by Region, 2017-18



Notes: Epanechnikov kernel used for univariate kernel density estimation. Per-pupil state revenues are restricted to <$30K to remove outliers.

Figure S1.5. District Per-Pupil Revenue – Local, Overall & by Region, 2017-18



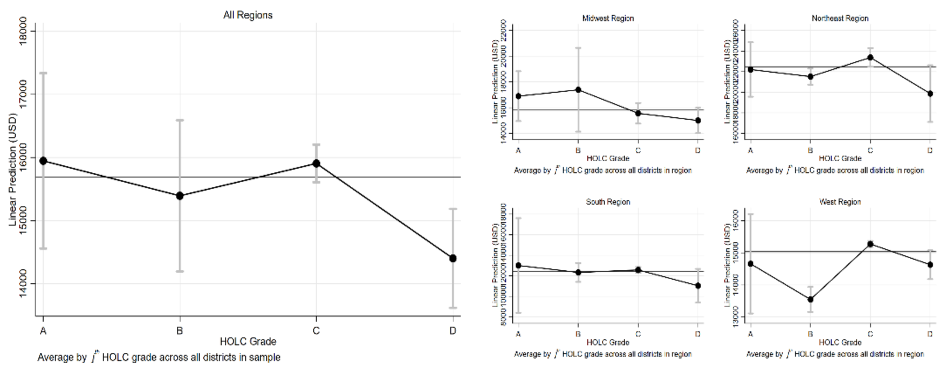
Notes: Epanechnikov kernel used for univariate kernel density estimation. Per-pupil local revenues are restricted to <$40K to remove outliers.

Table S1.3 Models of District Per-Pupil Revenues on 1935-1940 HOLC A-D Grades, USD 2018

	(1)				**(2)**			
	Total	*Federal*	*State*	*Local*	*Total*	*Federal*	*State*	*Local*
A	2,486	-431***	-3,016***	5,933***	1,546*	-722***	-3,670***	5,937***
	(1,896)	(146)	(765)	(1,671)	(822)	(139)	(752)	(1,228)
B	3,210***	-194**	-1,278*	4,681***	992*	-391***	-3,210***	4,592***
	(746)	(89)	(708)	(948)	(536)	(87)	(776)	(892)
C	982	294***	-243	930**	1,504***	253***	-461	1,713***
	(623)	(67)	(535)	(423)	(514)	(76)	(495)	(468)
Constant	14,453***	1,136***	7,514***	5,803***	14,402***	1,201***	7,960***	5,240***
	(815)	(56)	(532)	(462)	(396)	(62)	(438)	(425)
FE	N	N	N	N	Y	Y	Y	Y
R^2	0.029	0.113	0.023	0.105	0.643	0.324	0.515	0.493
F	6.532	22.21	6.314	10.12	4.076	20.16	8.754	11.56
N	1760	1760	1760	1760	1760	1760	1760	1760

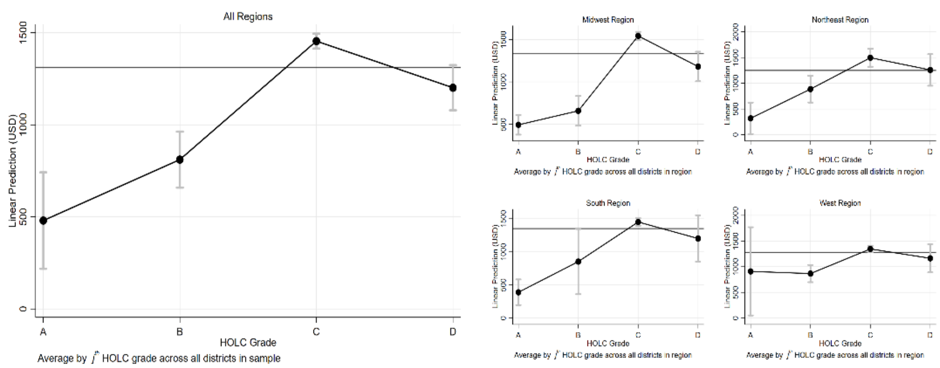
Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level
Note 2: For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. Models (1) and (2) use the NCES total, federal, state and local per-pupil revenue covariates. All regressions are weighted by 2017-18 district-level enrollment. Model (1) does not use fixed effects whereas Model (2) uses city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% levels*

Figure S1.6. District Per-Pupil Revenue – Total, Overall & by Region, 2017-18



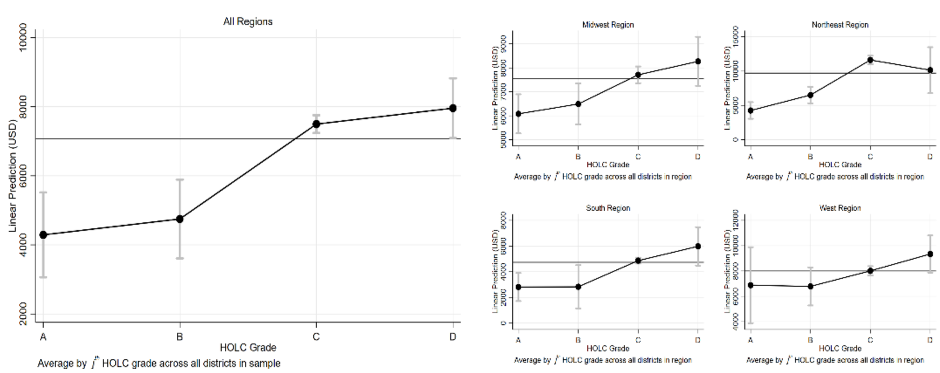
Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2017-18 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2017-18 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.

Figure S1.7. District Per-Pupil Revenue – Federal, Overall & by Region, 2017-18



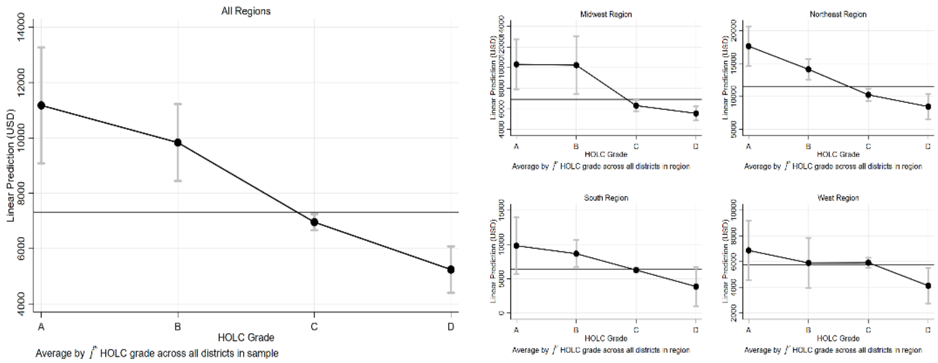
Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2017-18 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2017-18 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.

Figure S1.8. District Per-Pupil Revenue – State, Overall & by Region, 2017-18



Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2017-18 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2017-18 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.

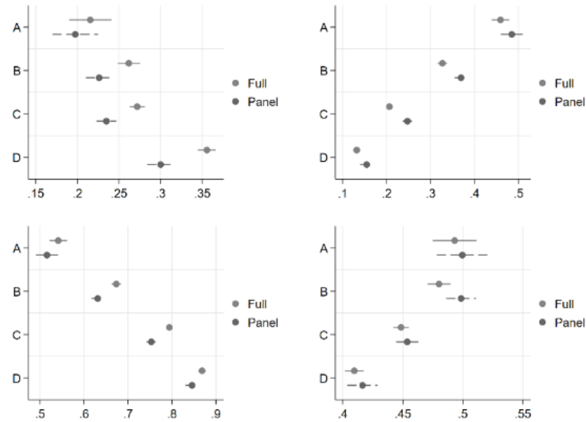
Figure S1.9. District Per-Pupil Revenue – Local, Overall & by Region, 2017-18 [CBSA FE]



Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2017-18 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2017-18 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.

Appendix S2. School Diversity Outcomes

Figure S2.1. HOLC A-D Coefficient Plots, Full vs. Panel Sample, 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 = 10,010 schools. The panel sample is the time series sample spanning 1988-89 through 2018-19 school years and includes np = 4,693 schools.

Table S2.1. HOLC A-D Means, SEs and Differences, Full vs. Panel Sample, 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 Diversity (1-D)	A	0.49	0.50	-0.006	0.329
		[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]		

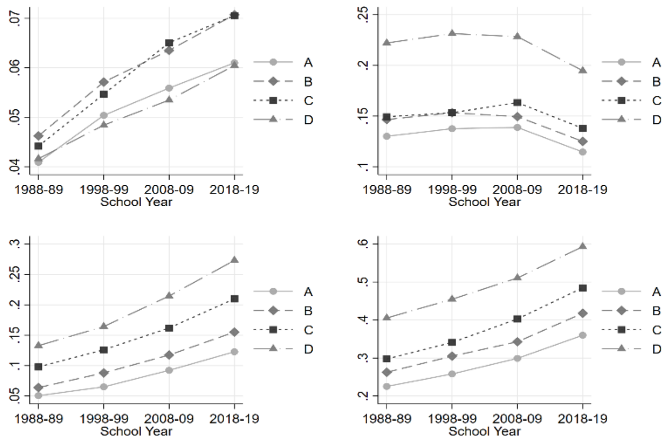
Notes: Means are HOLC A-D weighted averages from regressions of respective diversity outcomes on HOLC B-D indicators. Standard errors are Huber-White heteroskedasticity-consistent and are in brackets. Regressions are run using both the 2017-18 full sample and the 2017-18 data from the panel sample. The full sample is the original 2017-18 cross-sectional sample and includes all n = 10,010 schools. The panel sample is the time series sample spanning 1988-89 through 2018-19 school years and includes np = 4,693 schools.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.*

Table S2.2. HOLC A-D Averages Over Time, Full and Panel Samples, Exposure Index, 1988-2019

		HOLC Security Rating				
		Overall	**A**	**B**	**C**	**D**
White-Asian [a]	1988-89	4.4%	4.1%	4.6%	4.4%	4.2%
	1998-99	5.4%	5.0%	5.7%	5.5%	4.8%
	2008-09	6.2%	5.6%	6.4%	6.5%	5.3%
	2018-19	6.8%	6.1%	7.1%	7.1%	6.0%
	2018-19 [b]	7.5%	6.5%	7.6%	8.0%	7.3%
White-Black [a]	1988-89	15.9%	13.0%	14.7%	14.9%	22.2%
	1998-99	16.3%	13.8%	15.3%	15.3%	23.1%
	2008-09	16.4%	13.9%	14.9%	16.3%	22.8%
	2018-19	13.8%	11.5%	12.5%	13.8%	19.5%
	2018-19 [b]	15.0%	11.5%	13.3%	14.9%	20.1%
White-Hispanic [a]	1988-89	8.9%	5.1%	6.4%	9.8%	13.3%
	1998-99	11.3%	6.5%	8.8%	12.6%	16.4%
	2008-09	14.5%	9.2%	11.7%	16.2%	21.4%
	2018-19	18.8%	12.3%	15.5%	21.0%	27.3%
	2018-19 [b]	19.9%	12.8%	16.3%	22.0%	25.8%
White-Non-White [a]	1988-89	29.8%	22.5%	26.3%	29.8%	40.5%
	1998-99	33.7%	25.8%	30.5%	34.1%	45.5%
	2008-09	38.4%	29.9%	34.3%	40.3%	51.1%
	2018-19	45.9%	36.0%	41.8%	48.5%	59.3%
	2018-19 [b]	48.7%	36.9%	43.6%	51.1%	59.3%

Notes: a – Signifies the student race pair used in the exposure index calculation. For example, White-Black provides the White-Black exposure index, while factoring in all other student race groups. Exposure index calculations do vary if all other race groups apart from the relevant race pair are ignored; b – Exposure index values derived using the full 2018-19 cross-sectional data set. All other within-group (i.e., White-Black) values across time derived from the 1988-89 through 2018-19 panel data set.

Figure S2.2. HOLC A-D Averages Over Time, Exposure Index (1988-2018)



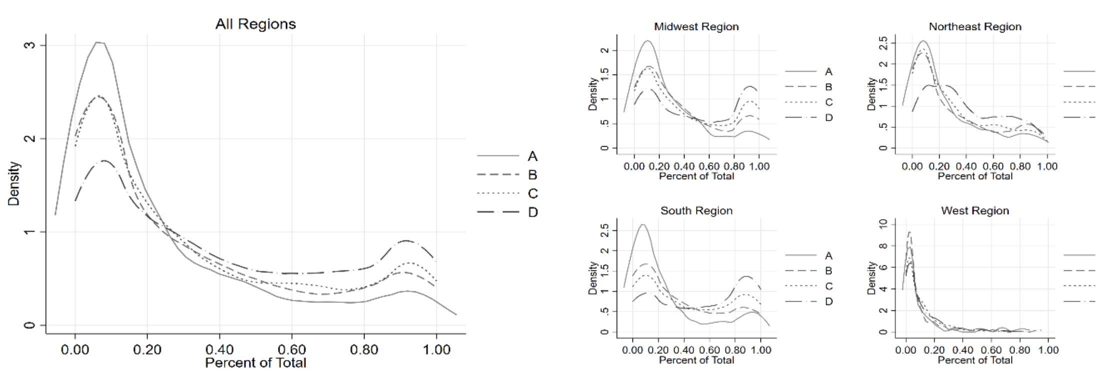
Notes: [top left] White-Asian Exposure Index; [top right] White-Black Exposure Index; [bottom left] White-Hispanic Exposure Index; [bottom right] White-Non-White Exposure Index. Exposure index calculations for each race pair factor in all other student race groups.

Table S2.3 Kolmogorov-Smirnov Equality of Distribution Test, Diversity Outcomes

		U.S. Census Bureau Region[b]				
		Overall	**Midwest**	**Northeast**	**South**	**West**
% Black	A vs. D [a]	0.26***	0.34***	0.32***	0.43***	0.17**
	B vs. D	0.16***	0.21***	0.25***	0.29***	0.12***
	C vs. D	0.13***	0.13***	0.2***	0.17***	0.03
% White	D vs. A	-0.54***	-0.54***	-0.58***	-0.44***	-0.65***
	D vs. B	-0.34***	-0.34***	-0.36***	-0.23***	-0.47***
	D vs. C	-0.16***	-0.17***	-0.19***	-0.10***	-0.19***
% Non-White	A vs. D	0.54***	0.54***	0.58***	0.44***	0.65***
	B vs. D	0.34***	0.34***	0.36***	0.23***	0.47***
	C vs. D	0.16***	0.17***	0.19***	0.10***	0.19***
Simpson's Diversity (1-D)	D vs. A	-0.18***	-0.26***	-0.18***	-0.22***	-0.41***
	D vs. B	-0.12***	-0.17***	-0.10***	-0.14***	-0.31***
	D vs. C	-0.09***	-0.12***	-0.11***	-0.08**	-0.17***

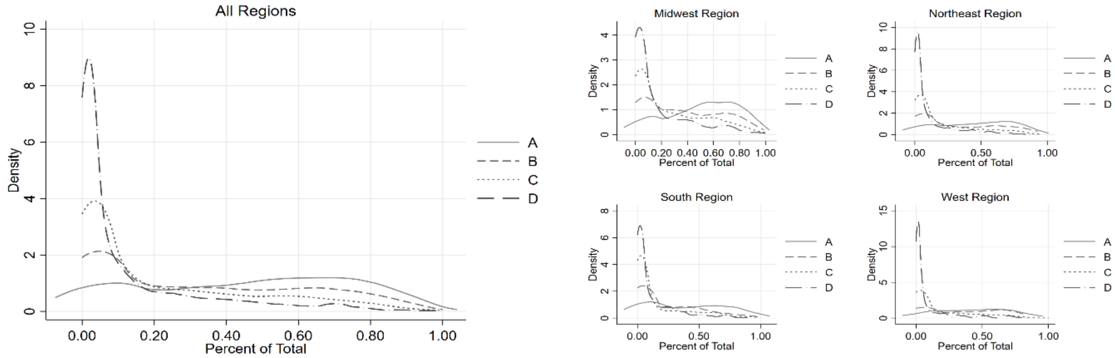
Notes: a – Tests the null hypothesis that the first HOLC grade distribution contains smaller values than the second HOLC grade distribution. For example, D vs. A test the null hypothesis that D distribution contains smaller values than A distribution; b – The values seen here, both for overall and each region, represent the largest difference between distributions in the direction as specified from the "Group" row input.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Figure S2.3. School Student Body % Black, Overall & by Region, 2018-2019



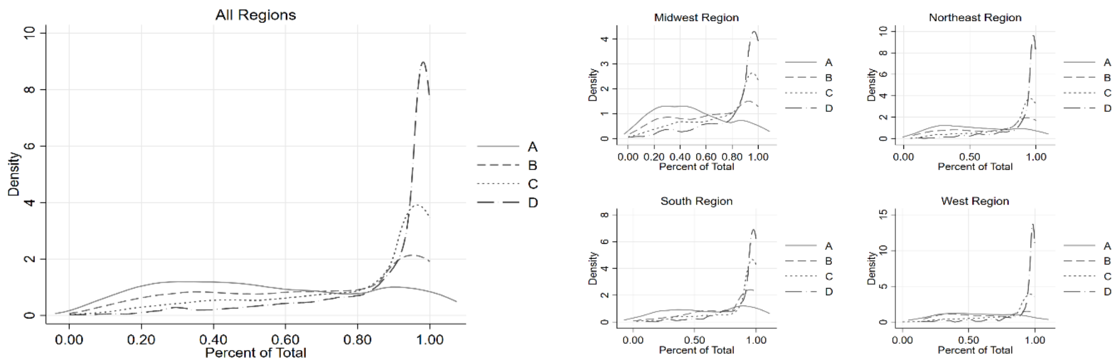
Notes: Epanechnikov kernel used for univariate kernel density estimation

Figure S2.4. School Student Body % White, Overall & by Region, 2018-2019



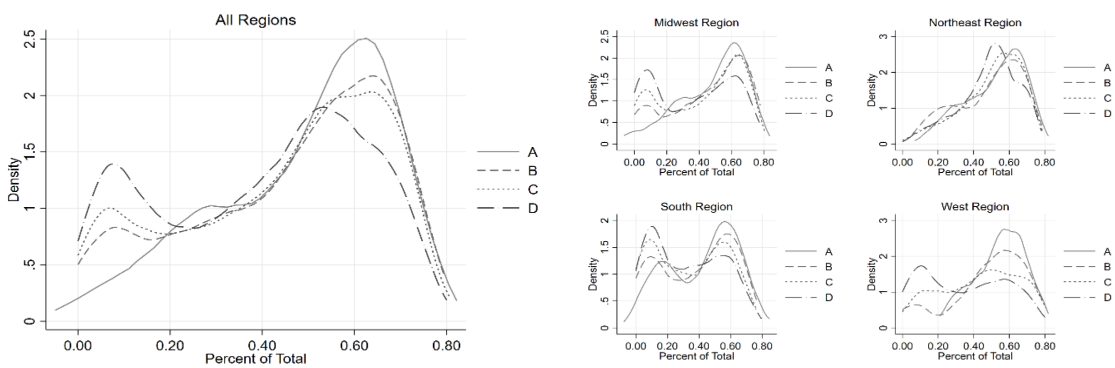
Notes: Epanechnikov kernel used for univariate kernel density estimation

Figure S2.5. School Student Body % Non-White, Overall & by Region, 2018-2019



Notes: Epanechnikov kernel used for univariate kernel density estimation

Figure S2.6. School Student Diversity, Overall & by Region, 2018-2019



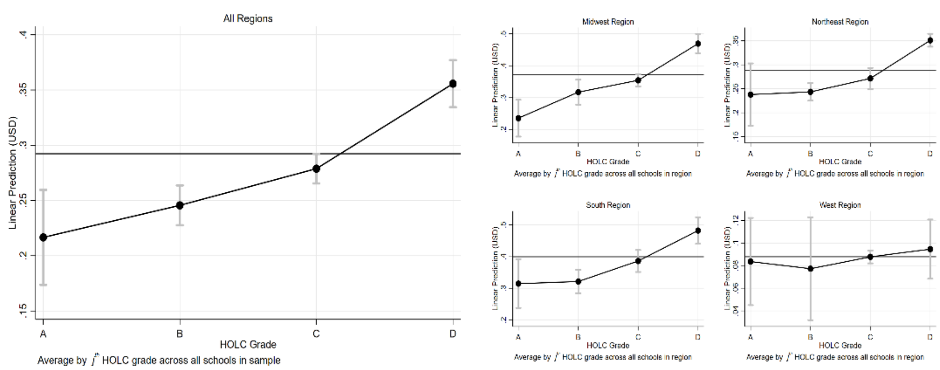
Notes: Epanechnikov kernel used for univariate kernel density estimation

Table S2.4 Models of School Demographics and Diversity on 1935-1940 HOLC A-D Grades, 2018-19

	(1)				**(2)**			
	% Black	*% White*	*% Non-White*	*Simpson's Diversity*	*% Black*	*% White*	*% Non-White*	*Simpson's Diversity*
A	-0.14***	0.33***	-0.33***	0.08***	-0.14***	0.30***	-0.30***	0.07***
	(0.03)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
B	-0.09***	0.20***	-0.20***	0.07**	-0.11***	0.18***	-0.18***	0.05***
	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)
C	-0.08***	0.07***	-0.07***	0.04***	-0.08***	0.06***	-0.06***	0.03***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Constant	0.36***	0.13***	0.87***	0.41***	0.36***	0.14***	0.86***	0.42***
	(0.04)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
FE	N	N	N	N	Y	Y	Y	Y
R^2	0.021	0.123	0.123	0.016	0.333	0.336	0.336	0.286
F	12.22	75.24	75.24	3.566	13.74	86.13	86.13	3.804
Observations	9709	9709	9709	9709	9709	9709	9709	9709

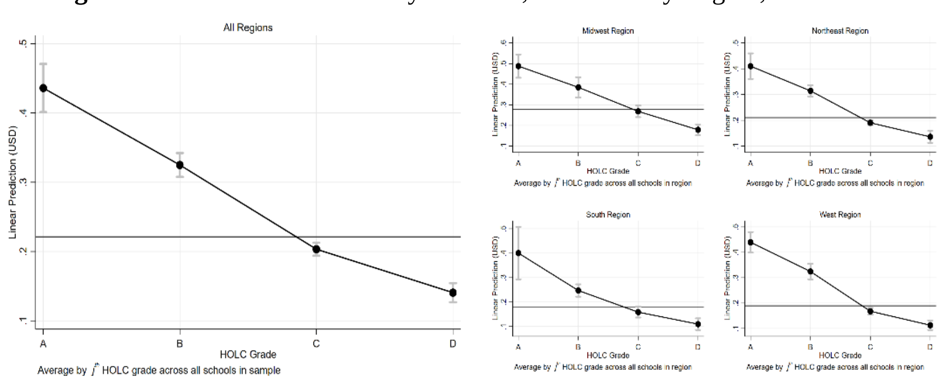
Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level
Note 2: For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. Models (1) and (2) use percent Black, White, Non-White and Simpson's Diversity (1-D) covariates. All regressions are weighted by 2018-19 school-level enrollment. Model (1) does not use fixed effects whereas Model (2) uses city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.*

Figure S2.7. School Student Body % Black, Overall & by Region, 2018-2019



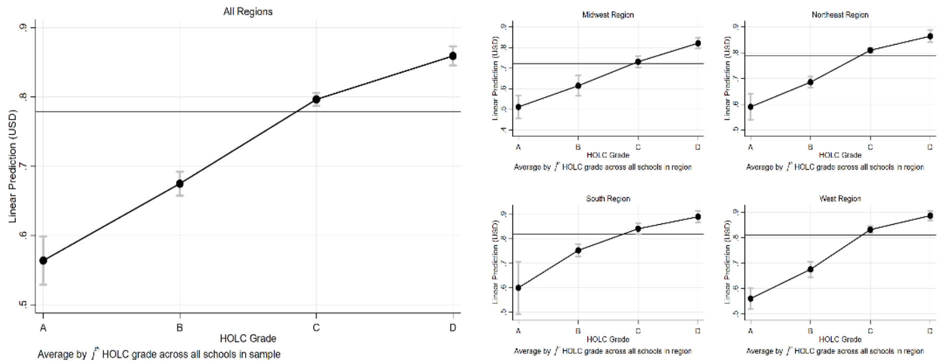
Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2018-19 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2018-19 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.

Figure S2.8. School Student Body % White, Overall & by Region, 2018-2019



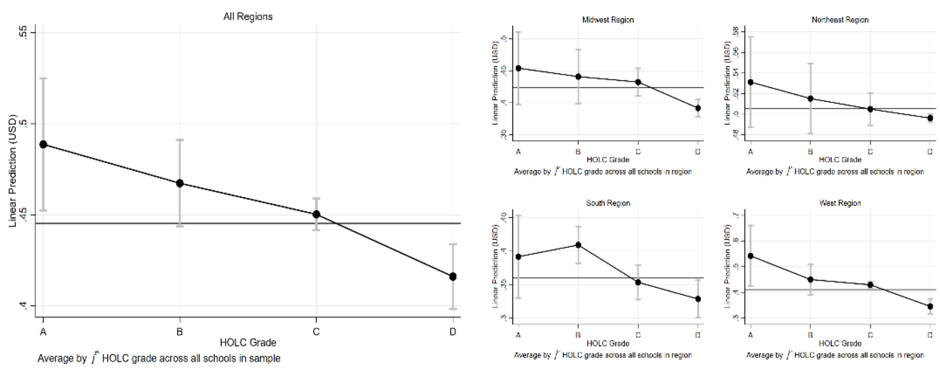
Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2018-19 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D *weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2018-19 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.*

Figure S2.9. School Student Body % Non-White, Overall & by Region, 2018-2019



Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2018-19 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2018-19 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.

Figure S2.10. Simpson's Diversity (1-D), Overall & by Region, 2018-2019



Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by 2018-19 student enrollment. Each black dot represents the weighted average outcome by HOLC A-D grade. The gray bands represent 95% confidence intervals. HOLC A-D weighted averages are derived from a city-level fixed effects regression of the above outcome on HOLC A-C indicators with 2018-19 student enrollment used as analytic weights. No controls are included. Standard errors are cluster-robust with clustering at the city-level.

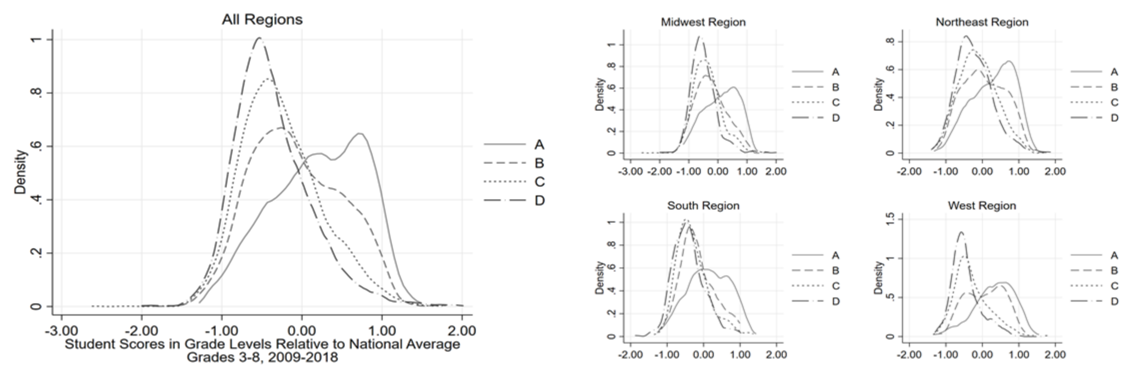
Appendix S3. School Performance Outcomes

Table S3.1. Kolmogorov-Smirnov Equality of Distribution Test, Student Performance Outcomes

		U.S. Census Bureau Region[b]				
		Overall	**Midwest**	**Northeast**	**South**	**West**
Avg. Student Math & ELA Scores	D vs. A[a]	-0.47***	-0.52***	-0.44***	-0.49***	-0.67***
	D vs. B	-0.25***	-0.25***	-0.22***	-0.24***	-0.45***
	D vs. C	-0.12***	-0.16***	-0.11***	-0.12***	-0.18***
Student Learning (Annual)	D vs. A	-0.07	-0.06	-0.11	-0.07	-0.12
	D vs. B	-0.05*	-0.02	-0.09**	-0.07	-0.12*
	D vs. C	-0.03	-0.02	-0.05	-0.08	-0.03
Avg. Student Test Score Change	D vs. A	-0.08**	-0.15**	-0.14**	-0.1	-0.1
	D vs. B	-0.05**	-0.07	-0.05	-0.08	-0.13**
	D vs. C	-0.04**	-0.07**	-0.06**	-0.06	-0.03

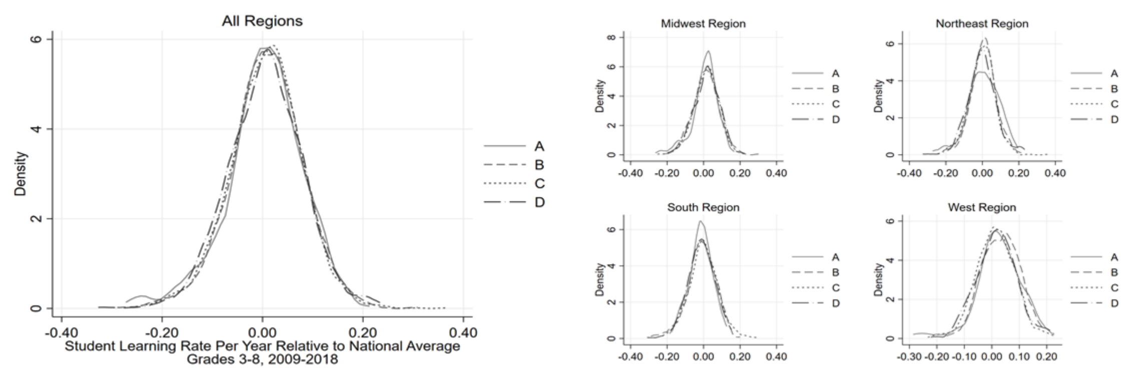
Notes: a – Tests the null hypothesis that the first HOLC grade distribution contains smaller values than the second HOLC grade distribution. For example, D vs. A test the null hypothesis that D distribution contains smaller values than A distribution; b – The values seen here, both for overall and each region, represent the largest difference between distributions in the direction as specified from the "Group" row input.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level*

Figure S3.1. School-Level Average Student Math & ELA Scores Relative to U.S. National Average, Grades 3-8, 2009-2018 (SD Units)



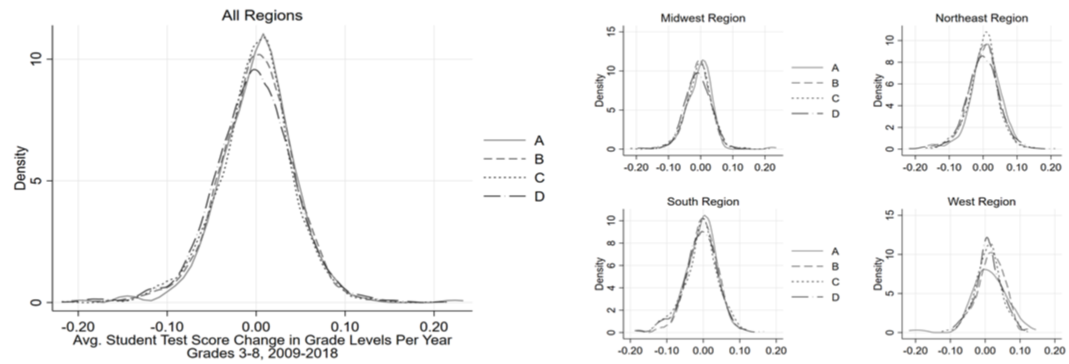
Notes: Epanechnikov kernel used for univariate kernel density estimation. SEDA OLS estimation used above. Results are not sensitive to Empirical Bayes estimation.

Figure S3.2. School-Level Student Learning Per Year Relative to U.S. National Average, Grades 3-8, 2009-2018 (SD Units)



Notes: Epanechnikov kernel used for univariate kernel density estimation. SEDA OLS estimation used above. Results are not sensitive to Empirical Bayes estimation.

Figure S3.3. School-Level Average Student Test Score Change, Grades 3-8, 2009-2018 (SD Units)



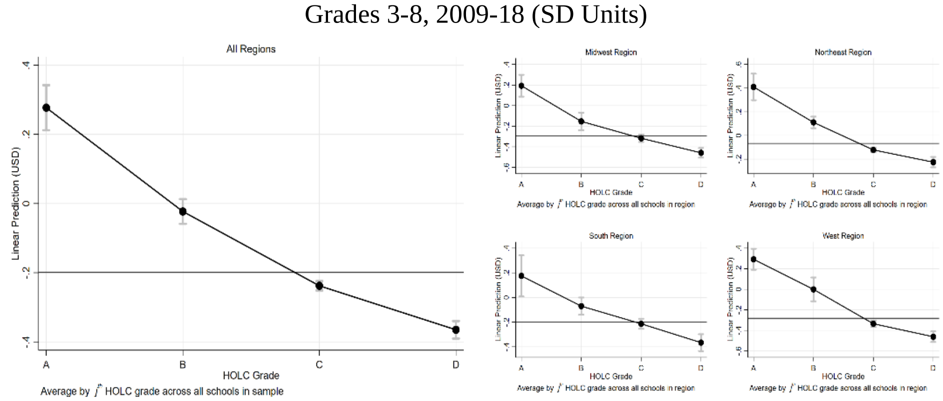
Notes: Epanechnikov kernel used for univariate kernel density estimation. SEDA OLS estimation used above. Results are not sensitive to Empirical Bayes estimation.

Table S3.2. Models of School Student Performance Measures on 1935-1940 HOLC A-D Grades, 2009-2018

		(1)			**(2)**	
	Avg. Test Scores	*Avg. Student Learning*	*Avg. Student Growth*	*Avg. Test Scores*	*Avg. Student Learning*	*Avg. Student Growth*
A	0.571***	-0.003	0.005*	0.643***	0.006	0.010***
	(0.048)	(0.007)	(0.003)	(0.040)	(0.006)	(0.003)
B	0.306***	0.002	0.002	0.342***	0.008*	0.007***
	(0.037)	(0.006)	(0.003)	(0.028)	(0.004)	(0.002)
C	0.119***	0.005*	0.003**	0.127***	0.005*	0.004***
	(0.016)	(0.003)	(0.001)	(0.015)	(0.003)	(0.001)
Constant	-0.350***	0.009*	-0.002	-0.365***	0.007***	-0.003***
	(0.057)	(0.005)	(0.004)	(0.013)	(0.002)	(0.001)
FE	N	N	N	Y	Y	Y
R^2	0.084	0.001	0.001	0.266	0.142	0.194
F	51.44	1.294	1.624	91.85	1.510	5.100
N	6,890	5,193	6,183	6,890	5,193	6,183

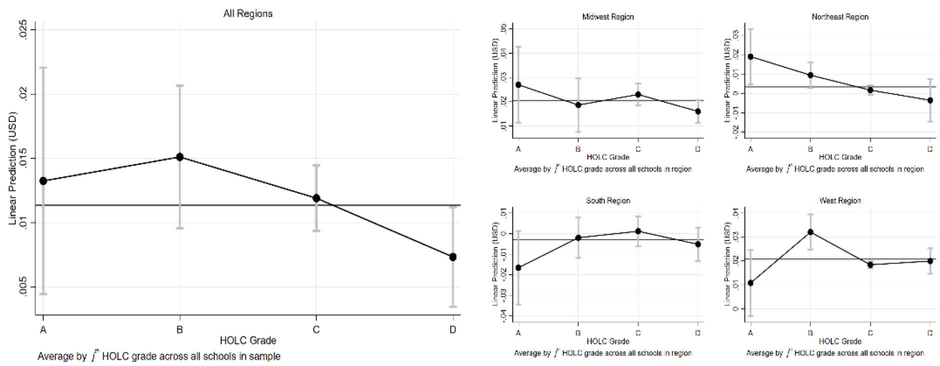
Note 1: Cluster-robust standard errors are in parentheses with clustering done at the city-level.
Note 2: For each model we regress the outcome on HOLC grade indicators with the "D" security rating as the reference category. No controls are included. Models (1) and (2) use the same set of SEDA school-level student performance covariates. All regressions are weighted by 2009-2018 total math and ELA tests for pooled SEDA estimates. Results are not sensitive to alternative weighting by 2017-18 total students. Model (1) does not use fixed effects whereas Model (2) uses city-level fixed effects to account for any differences that are fixed at the local level and that differ between CBSAs. All outcomes are in terms of standard deviation (SD) units where 1/3 of a SD unit approximates one grade-level.
**** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% levels*

Figure S3.4. Average Student Math & ELA Scores – Overall & by Region, Relative to U.S. National Average, Grades 3-8, 2009-18 (SD Units)



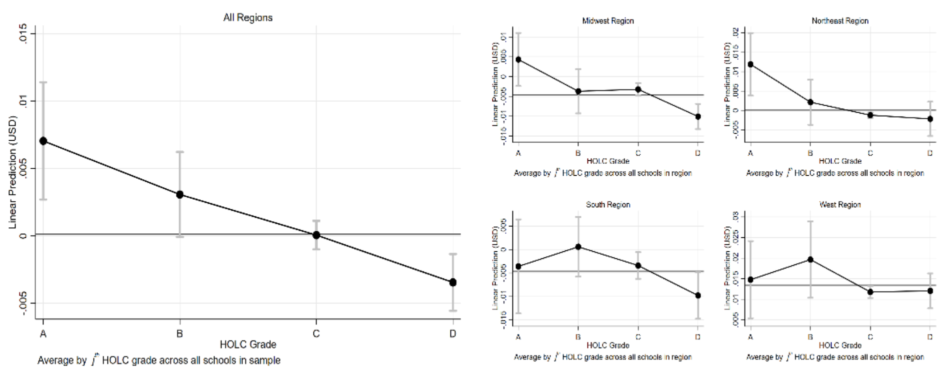
Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by total number of math and ELA tests for pooled SEDA estimates. Each black dot represents the weighted average outcome by HOLC A-D grade weighted by total number of math and ELA tests, and are derived from a regression of each respective outcome on HOLC A-C indicators with fixed effects at the city-level. Standard errors are clustered at the city level.

Figure S3.5. Average Student Learning – Overall & by Region, Relative to U.S. National Average, Grades 3-8, 2009-18 (SD Units)



Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by total number of math and ELA tests for pooled SEDA estimates. Each black dot represents the weighted average outcome by HOLC A-D grade weighted by total number of math and ELA tests, and are derived from a regression of each respective outcome on HOLC A-C indicators with fixed effects at the city-level. Standard errors are clustered at the city level.

Figure S3.6. Average Student Test Score Change – Overall & by Region, Relative to U.S. National Average, Grades 3-8, 2009-18 (SD Units)



Notes: For each plot, both overall and by region, the horizontal line represents the grand mean, weighted by total number of math and ELA tests for pooled SEDA estimates. Each black dot represents the weighted average outcome by HOLC A-D grade weighted by total number of math and ELA tests, and are derived from a regression of each respective outcome on HOLC A-C indicators with fixed effects at the city-level. Standard errors are clustered at the city level.