Federal Reserve Bank of Chicago

The Long-run Effects of the 1930s Redlining Maps on Children

Daniel Aaronson, Bhashkar Mazumder, Daniel Hartley, and Martha Stinson

April 2022

WP 2022-13

https://doi.org/10.21033/wp-2022-13

The Long-run Effects of the 1930s Redlining Maps on Children*

Daniel Aaronson
Federal Reserve Bank of Chicago
daaronson@frbchi.org

Bhashkar Mazumder
Federal Reserve Bank of Chicago
bhash.mazumder@gmail.com

Daniel Hartley
Federal Reserve Bank of Chicago
daniel.a.hartley@chi.frb.org

Martha Stinson
U.S. Census Bureau
Martha.Stinson@census.gov

April 2022

Abstract: We estimate the long-run effects of the 1930s Home Owners Loan Corporation (HOLC) redlining maps by linking children in the full count 1940 Census to 1) the universe of IRS tax data in 1974 and 1979 and 2) the long form 2000 Census. We use two identification strategies to estimate the potential long-run effects of differential access to credit along HOLC boundaries. The first strategy compares cross-boundary differences along HOLC boundaries to a comparison group of boundaries that had statistically similar pre-existing differences as the actual boundaries. A second approach only uses boundaries that were least likely to have been chosen by the HOLC based on our statistical model. We find that children living on the lower-graded side of HOLC boundaries had significantly lower levels of educational attainment, reduced income in adulthood, and lived in neighborhoods during adulthood characterized by lower educational attainment, higher poverty rates, and higher rates of single-headed households.

*We thank Adriana Lleras-Muney and Shari Eli for many helpful comments. We also thank Raven Molloy, Trevon Logan and Sun Kyoung Lee for reading and discussing our paper at conferences. We thank participants at the Federal Reserve System Committee on Regional Analysis Conference, the NBER Economics of Mobility Workshop and the Race, Ethnicity, Gender, and Economic Justice Virtual Symposium held at Yale University. The views expressed are those of the authors and do not necessarily represent the views of the U.S. Census Bureau, the Federal Reserve Bank of Chicago, the Board of Governors of the Federal Reserve System, or its staff. This research was conducted as part of an Internal Census Project. All results using internal data have been reviewed to ensure that no confidential information is disclosed (DRB clearance numbers CBDRB-FY22-CES17-001 and CBDRB-FY22-115).

1. Introduction

There is accumulating evidence that neighborhood exposure during childhood has long-run effects on an individual's socioeconomic status later in life, such as their educational attainment, income, and upward income mobility (e.g. Chetty, Hendren and Katz, 2016, Chyn, 2018, Chetty and Hendren, 2018). What is less clear is what it is about neighborhoods that leads to long-run effects, a point emphasized in Chyn and Katz's (2021, p. 25) recent review of the neighborhood effects literature: "Future work related to mechanisms remains an ongoing research issue. For example, we know little about the relative importance of the different mechanisms that are typically 'bundled' together within a neighborhood."

We focus on a historical episode that may shed light on one particular aspect of neighborhoods that can impact children's later life outcomes, namely access to credit in the community you are raised in. Specifically, we examine the effects of the so-called "redlining maps" that were drawn in the mid- to late-1930s by a federal agency called the Home Owners Loan Corporation (HOLC). The HOLC played a crucial role in stabilizing housing markets by buying and refinancing mortgages during the foreclosure crisis of the first half of the 1930s. In 1935, they were tasked with creating a new housing appraisal system that involved producing color-coded maps for over 200 cities to categorize the riskiness of lending to households in different neighborhoods. Neighborhoods received one of four letter grades (A to D) that corresponded to their estimated risk. While the HOLC used a large set of housing and economic measures, many collected by their surveyors, to categorize neighborhood credit risk, they also incorporated demographic information such as the racial, ethnic, and immigrant composition of neighborhoods into these creditworthiness profiles.

There is an ongoing debate about the extent to which the HOLC maps influenced private lending activity and therefore could have had an impact on residents. Although the maps were supposed to be confidential (e.g. Hillier, 2003, Fishback et al. 2021), there are known examples of breaches in confidentiality (e.g. Michney, 2021), suggesting that policies to prevent their dissemination were not always strictly adhered to. In addition, many private real estate and banking professionals throughout the country were involved in producing or reviewing the maps, thereby providing access to their underlying content which they could have retained for future use or shared with others. Some have also argued that the HOLC had a broad influence on the appraisal process used by private lenders nationwide by creating new standardized practices that emphasized the characteristics of neighborhoods. For example, some banks created their own version of the HOLC maps. Finally, the maps were shared with the Federal Housing Administration (FHA), which drew their own appraisal maps shortly after the HOLC, and the HOLC maps could have influenced FHA activity. The FHA played a critical role in the housing market by deciding whether to insure mortgages and were notorious for following discriminatory practices (Fishback et al 2021). Each of these potential avenues raise the prospect that the content of the HOLC maps, along with the broader use of a government-backed appraisal system, may have either directly or indirectly influenced access to credit for households in many urban neighborhoods. Although historical research on these different channels remains active, and we discuss some of this work further below, we may never fully know which of these mechanisms were the most pertinent.

Several studies have used the HOLC maps along with various research strategies to estimate their long-run effects on neighborhoods. In our own work, we find that the maps affected the trajectory of neighborhoods, as well as US cities, over the next 80 years: increasing racial segregation while reducing home ownership, house values, and rents (Aaronson, Hartley, and Mazumder 2021). We also document significant effects on *aggregate* socioeconomic outcomes for children living in these neighborhoods late in the 20th century using the Chetty et al (2018) Opportunity Atlas data (Aaronson, Faber, Hartley, Mazumder and Sharkey 2021). Other studies examine either causal effects or descriptive associations of the HOLC maps with a wide range of neighborhood level outcomes, such as housing, crime, firearm violence, schooling, birth outcomes, and the availability of greenspace.[1]

In this paper, we expand upon this literature by asking what happened to the later life educational attainment and income of children who grew up in differentially graded neighborhoods *at the time when*

[1] These include Appel and Nickerson (2016), Krimmel (2018), Anders (2021), Jacoby et al. (2018), Nardone et al. (2020), Hoffman et al (2020), Lukes and Cleveland (2021), and Hynsjö and Perdoni (2022).

the maps were first drawn in the late 1930s. To undertake this exercise, we assemble a unique dataset that links the 1940 complete count Census with 1) the universe of tax records in 1974 and 1979 and 2) the long form (1 in 6) sample of the 2000 Census. This linked data allows us to make several novel contributions. First, we are able to better isolate the direct effects of credit access by focusing on the initial treated cohorts before any long-run changes to impacted neighborhoods fully materialized. In contrast, those who were affected in later decades were exposed to a set of accumulated place effects, such as those that occurred through the resorting of individuals across neighborhoods or possibly through many urban policies that were more likely to occur in redlined areas. Second, in contrast to the existing literature which has focused exclusively on aggregate neighborhood level outcomes, we are able to look at *individual* outcomes. This is important because we can track people regardless of where they moved later in life. Our approach therefore more closely resembles the modern literature that uses experimental or quasi-experimental methods to trace the impact of specific policies that change neighborhood exposure during childhood. One important point to emphasize is that by focusing on a sample of children living in these neighborhoods during the 1930s, the vast majority of our sample consists of White children.

Our methodology closely follows Aaronson, Hartley, and Mazumder (2021), which we refer to as "AHM." Like AHM, we focus our analysis on very narrow areas (boundary buffer zones) surrounding borders demarcating C neighborhoods from D neighborhoods (D-C) or B neighborhoods from C neighborhoods (C-B). Using these buffer zones reduces, but does not eliminate, confounding factors that caused differences across neighborhoods with different grades. To identify causal effects, we then follow two strategies. First, we compare the cross-boundary difference along actual "treated" HOLC boundaries to a comparison group of cross-boundary differences from boundaries that could have been drawn but were not. Critically, these comparison boundary differences are reweighted so that they are statistically similar to the treated cross-boundary differences in observable characteristics and trends in these characteristics. AHM show that this additional contrast eliminates pre-trends, which provides indirect support for the validity of the empirical approach. They also show that the difference between the treated and comparison groups emerged only in subsequent decades after the maps were drawn.

Our second approach uses a statistical model to identify the subset of treated boundaries that were least likely to have been chosen by the HOLC. The intuition is that such boundaries may have been more idiosyncratically drawn. For example, they may have been chosen somewhat arbitrarily to close a polygon simply to define a neighborhood on a map. AHM show that there are no pre-existing cross-boundary differences or trends for this subset of boundaries. This approach is also appealing because there is no need to create a comparison group.[2]

Using either approach, we find significant effects of being on the lower-graded side of a D-C boundary (what is often referred to as "redlining") or C-B boundary (what we refer to as "yellowlining") on the educational attainment and adult income of children living in these neighborhoods in 1940. Specifically, we find that living on the lower-graded side leads to a decrease in years of education of about 0.2 years regardless of the border type (D-C or C-B). For annual adjusted gross income, we find large negative effects that are statistically significant but range in magnitude from $1,500 to $6,000 (in 2000$), depending on the border type and the estimation strategy used. Averaging across our estimates, we find that growing up on the lower-graded side of a boundary lowers annual real wage and salary income by 2.9 percent and annual real adjusted gross income by 4.3 percent relative to the mean levels in these areas.

The effects on income appear to be especially large along C-B boundaries. AHM similarly found larger and more persistent effects along these boundaries for some neighborhood-level outcomes such as house values. The negative effects on income appear to be especially strong for individuals with capital income (e.g. financial or real estate investments) who complete a schedule D form and are somewhat larger for children whose parents were homeowners in 1940. However, we generally did not find meaningful

[2] Aaronson, Hartley, and Mazumder (2021), Anders (2021), and Hynsjö and Perdoni (2022) also use an identification strategy based on the population cutoff of 40,000 that determined which cities were redlined. We did not pursue that strategy here.

income differences when we stratified the sample by several other characteristics available in the 1940 Census, including race and gender. That said, our sample of Black families is relatively small, as the vast majority of children who grew up in urban neighborhoods at the time were White.

We also take advantage of the geographic location of our sample of 1940 children at the time they filed their taxes in the 1970s to measure the socioeconomic characteristics of their neighbors in adulthood. Specifically, we construct an index based on six community characteristics that include measures of education, family structure, labor markets, and poverty. We find that children living on the lower-graded side of a D-C or C-B boundary in 1940, on average, live in a neighborhood over three decades later that is 0.02 to 0.04 standard deviation lower on this index. Much of this effect is driven by the educational attainment levels in these neighborhoods, although the poverty rate and the rate of single headed households also plays a role in some of our estimates. We also directly examined the likelihood of migration and find economically small negative effects on geographic mobility; indeed, nearly our entire sample left their childhood neighborhood by the 1970s.

An important issue is how to interpret our findings considering the paucity of historical information on the use of the maps. As noted earlier, private banks and the FHA drew their own color-coded maps that may have been influenced by the HOLC maps, which were supposed to have been confidential and limited in distribution. Fishback et al (2021) point out that the FHA's discriminatory practices predated the HOLC maps and that the FHA continually updated their maps (which largely no longer exist) over time suggesting that the influence of the HOLC maps may have been minimal. They also find little change in FHA loan insurance provision in three cities during the period in which the HOLC maps were made available to them. We discuss this issue in greater detail in the next section. If these other maps had causal effects, then our estimates may be picking up some combination of the direct effects of the HOLC maps, along with any overlap from other private or public maps that shared common neighborhood grades and borders with the HOLC maps. In that case, our estimates should be interpreted as a proxy for the overall effect of the general practice of redlining/yellowlining that arose largely through federal housing policies. On the other hand, this also implies that our estimates may actually be *understating* the full effects of discriminatory housing practices as measurement error could be introduced if some of the HOLC neighborhood grades or borders differed from those drawn in other, perhaps practically more impactful, maps from the FHA or private lenders.

2. Historical Background

This section provides a brief historical background on the Home Owners Loan Corporation (HOLC) maps.[3] The HOLC was created under the auspices of the Federal Home Loan Banking Board (FHLBB) in 1933 in order to help stabilize housing markets reeling from the Great Depression. Their specific early responsibility was to refinance troubled mortgages, and, by 1935, roughly 1 million had been modified. Previous studies have shown that the HOLC was not discriminatory in their own lending practices (Hillier, 2003, Fishback et al, 2021).[4]

Largely after this program was complete, the HOLC was tasked by the FHLBB to create a new appraisal system to help secure the future health of housing markets. As part of this effort, the HOLC drew color-coded residential "security' maps for 239 cities between 1935 and 1940, with grades assigned to neighborhoods between A and D based on their perceived level of riskiness. Neighborhoods that were considered the least risky were given an A grade and colored in green. The next three grades, B, C, and D were colored in blue, yellow and red, respectively. The HOLC used a large set of housing and economic measures that seem germane to categorize aggregate creditworthiness. But they also incorporated demographic information such as the racial, ethnic, and immigrant composition of neighborhoods. Woods (2012) argues that in addition to creating the maps, the HOLC played an important role in systemizing, nationalizing, and coordinating discriminatory practices through its new appraisal system. For example, regulatory policies required banks to create their own versions of the HOLC maps.

[3] A more complete description is in Aaronson, Hartley, and Mazumder (2021) and sources cited therein.
[4] To be clear, our current and previous analysis is not informed in any way by the loans directly financed by the HOLC through 1935 but rather is exclusively focused on the impact of the City Map program.

There is an ongoing debate among researchers about the extent to which the maps may have impacted access to credit either by influencing private lenders directly, or by affecting mortgage insurance provision decisions undertaken by the Federal Housing Administration (FHA). For example, Greer (2012) suggests that thousands of local real estate professionals were involved in the creation of the initial maps and that hundreds of officials at lending institutions in large cities may have been shown copies of the maps between 1938 and 1940 in the process of being interviewed by the HOLC.[5] AHM highlight one notable example where 8 of the 14 individuals who reviewed the maps for Cleveland were from local lending institutions or appraisers.[6] In addition, although, the HOLC maps were supposed to be held confidential and to not be shared outside the agency, this policy was not strictly adhered to and there are a few examples of breaches of confidentiality that have recently been discovered.[7]

Potentially more important, the maps could have been shared with the Federal Housing Administration (FHA), a separate Federal agency which drew their own neighborhood maps in order to decide whether to insure mortgages.[8] Michney (2021) highlights that the HOLC maps were an important input into the FHA's maps (p.19): "Crucially, the new evidence cited here almost certainly confirms Jackson's original contention that HOLC's maps factored into the FHA's own mapmaking process and quite plausibly that agency's racially-informed denial of access to its mortgage guarantees." Light (2020) also documents evidence that the HOLC maps impacted FHA mapmaking and policies and that the two

[5] Also see Jackson (1980).
[6] See https://library.osu.edu/documents/redlining-maps-ohio/area-descriptions/CuyahogaCounty_Explanation_and_A1-A31_Area_Description.pdf. On page 2 the document notes: "This map and the area descriptions have been prepared in collaboration with, and carefully checked by, competent local real estate brokers, appraisers, and mortgage lenders.". On page 4, it states: "The following men reviewed the entire map, materially aiding the completed work with their constructive suggestions and criticisms:" and then lists individuals including an Assistant Vice President in the loan department of the Cleveland Trust, a former President of the Real Estate Board, the President of the Ohio Savings and Loan, the President of the Cleveland chapter of the Society of Residential Appraisers, the Secretary of the Cuyahoga County Savings and Loan League, the President of the Lorain Street Bank, a manager in the mortgage loan department of National City Bank, a property sales manager for Cleveland Trust and several appraisers and city officials. This just happens to be one city for which we found such a list and it would be useful for future researchers to look for similar lists of the reviewers of the HOLC maps for other cities.
[7] Michney (2021) describes several examples of the HOLC maps being shared with private actors as well as with local and regional managers (p. 15-16), although he concludes that the maps were largely kept confidential from private actors.
[8] See Light (2010), Woods (2012), Hillier (2013), Michney (2021) and Fishback et al (2021).

agencies were closely connected.[9] Fishback et al (2021) point out that the HOLC maps could have been shared with the FHA as early as 1937 (p. 5).[10] They also note that, by 1942, the HOLC maps would have been less useful since the FHA were able to obtain block-level data from the 1940 Census. Even then, that leaves up to a five-year window in which the FHA could have incorporated the HOLC maps into their own.

Recent archival work by Fishback et al (2021) makes an important contribution in advancing the literature on FHA loan insurance activity by collecting and analyzing new data for three cities -- Baltimore, Peoria, and Greensboro -- over the period from 1935 to 1940. They show that the FHA was already practicing redlining policies as early as 1935 -- i.e. before the HOLC maps were drawn. Moreover, they find little change in loan insurance activity in these three cities through the late 1930s in response to the release of HOLC maps. However, their annual time series cannot rule out a timeline in which the FHA's loan insurance was potentially influenced by the HOLC maps to a modest degree.[11] Future large scale quantitative work collecting data covering FHA activity in many more cities may be needed to resolve this issue more definitively.

Regardless, our approach of focusing on the causal effects of the HOLC maps on children growing up in affected neighborhoods can be viewed as a proxy for the effects of federal housing policies that arose through some combination of the direct effect of the HOLC maps themselves on private lending decisions,

---

[9] Light (2020) writes in Footnote 85: "Ample evidence indicates the two agencies exchanged ideas. Coming together at one-time meetings such as the National Appraisal Forum, they maintained more regular communications thanks to the U.S. Central Housing Committee, which included representatives from both agencies on its Joint Committee on Appraisal and Mortgage Analysis. FHA records indicate the agency kept the HOLC security maps on file in connection with the construction of its Economic Data System (one of Hoyt's other projects) and comments from Federal Home Loan Bank Board general counsel Horace Russell on how the FHA "was fortunate in being able to avail itself of much of the training and experience in appraisal and the development of appraisal data by Home Owners Loan Corporation" underscores the two agencies'close ties." Furthermore, a 1938 FHA underwriting manual used examples that were taken directly from the HOLC (Woods, 2012).

[10] They write: "a 1942 document states that the HOLC shared copies of the entire set of maps with the FHA upon completion" (p. 10) and that this sharing could have taken place as early as 1937 (p. 10).

[11] The time series plot of FHA loan activity from 1935 to 1940 in Fishback et al (2021) (figure 4) shows that the level of mortgage loan insurance in HOLC D-graded areas was very small between 1935 and 1937 but becomes virtually non-existent after 1937. Thus, it is precisely after 1937, when the HOLC maps may have been shared with the FHA, that their lending activities appear to be most consistent with redlining. The maps available at the University of Richmond's Digital Scholarship lab, which may have been the final maps, are dated as follows: Baltimore, May 1, 1937; Greensboro, June 2, 1937, and Peoria, October 28, 1938.

and their indirect effects through FHA activity and maps drawn by individual private lenders. If it was the case that the HOLC maps truly had no direct effects on private lending and only affected neighborhoods through their partial influence on the FHA, then our estimates are likely an *underestimate* of the full effects of federal housing policies more broadly. For example, if the original maps created by the FHA that were used for their loan insurance decisions still existed and were available to researchers, then we could use a similar methodology to estimate the effects of the FHA policies.[12] In that case, relative to directly estimated effects of the FHA maps, our estimates based on using the HOLC maps as a proxy should produce attenuated estimates of the effects of the FHA maps due to the measurement error that would arise from using borders and security grades that did not always line up across the two maps.[13]

3. Data and Methodology

HOLC Maps

Our analysis uses the geocoded renderings of the original HOLC maps for the 149 cities used in AHM. The maps were made available by the University of Richmond Digital Scholarship Lab.[14] Residential neighborhoods in each city received a grade of A, B, C, or D and in some cases adjacent neighborhoods receive different grades, thereby creating boundaries or borders between differently-graded neighborhoods.[15] Our analysis will utilize these boundaries using two statistical approaches that we describe

[12] Fishback et al (2021) cite Greer (2014) and Sagalyn (1980) who suggest that the FHA maps were destroyed as the result of a lawsuit in 1969. We are aware of only two FHA maps that still exist, one for Chicago and for Washington D.C. However, there is no HOLC map for Washington D.C. Aaronson, Hartley and Mazumder (2021) report that 82 percent of Chicago has the same grade on both the HOLC and FHA maps when neighborhoods are weighed by population, including 86 percent overlap for D-graded areas. See Xu (2021) for a more detailed comparison of the grading differences between the two maps.

[13] Consistent with this possibility of attenuation due to measurement error, Xu (2021) finds meaningful discrepancies in the grading of neighborhoods in Chicago between the two agencies and estimates larger effects of the FHA maps than the HOLC maps on outcomes such as median home values. For example, Xu writes "14.4% of the FHA-graded D tracts were graded C by HOLC, whereas 26.3% of the FHA grade C tracts were graded D by HOLC". We should note, however, that the coarseness of the Chicago map makes the comparison of FHA and HOLC maps in Xu (2021) interesting but not necessarily representative of what would be found in other cities throughout the country. AHM find that the effects of redlining and yellowlining were fairly small in Chicago compared to other cities (see their Appendix Table A9).

[14] See Aaronson, Hartley, and Mazumder (2021) for more details on these cities. The geocoded maps were obtained at http://dsl.richmond.edu/panorama/redlining/

[15] We use the terms boundaries and borders interchangeably.

below. Specifically, we create an ID for each straight-line segment of an HOLC boundary that is at least a quarter mile long. We focus only on boundaries demarcating either a D neighborhood from a C neighborhood ("D-C") or a C neighborhood from a B neighborhood ("C-B").[16] We then create "boundary buffer zones" or "buffers" for short, which are rectangles that extend a quarter of a mile on each side of a boundary. Each boundary has two buffers – the lower-graded side (LGS) and higher-graded side (HGS).

Census Linked Data:

Our sample of children aged 16 and under come from the restricted version of the 1940 full-count Decennial Census (Minnesota Population Center and Ancestry.com 2013). We use the U.S. Census Bureau's Protected Identification Key (PIK) as a unique ID to match the 1940 Census children to the IRS Form 1040 tax returns from 1974 and 1979 and the 2000 Decennial Census long-form. The PIKs for the 1940 sample were created by matching the 1940 Census data to the SSA Numident file, a database of individuals who have ever applied for a Social Security Number (SSN). The match is largely based on demographics and geography.[17]

Of the 8.48 million children under age 16 who lived in an HOLC city, approximately 5.65 million of them were assigned PIKs, for a match rate of 67 percent. Individuals who failed to match to the Numident either had missing, inaccurate, or non-unique linkage keys or were old enough that they never applied for an SSN and so were not captured in the Numident database. While this latter issue presents significant challenges for matching older individuals, it has much less of an impact on our sample of children in 1940.

---

[16] A-graded neighborhoods represent the smallest share of neighborhoods in the maps and neighborhoods tend to border neighborhoods with either the same grade or one grade lower or higher. For these reasons the C-D and D-C boundaries are the most prevalent.

[17] PIKs are similar to SSNs but anonymized. Massey, Genadek, Alexander, Gardner, and O'Hara (2018) describe the system that was developed for matching the 1940 Census. For children living with their parents in 1940, the matching algorithm uses both the child's and parents' names, since Numident records contain both individual, mother, and father names. This is especially helpful for women who, at this time, commonly changed their last names at marriage. The father's name on the Numident record is the equivalent of the woman's birth surname and will often match to the woman's 1940 last name when she was a child. Other matching data fields include age on April 1, 1940, state or country of birth, sex, and state of residence in 1940 and 1935 which is known on the 1940 Census and can be partially inferred on the Numident based on the first 3 digits of the SSN.

A similar PIK assignment process was used for the 2000 Census, enabling its linkage to the 1940 Census (see Wagner and Lane (2014) for details). For this analysis, we focus on educational attainment,[18] namely total years of completed schooling as well as indicators for having completed high school or more, attended college or more, and having completed college or more.[19] This sample has a few drawbacks. First, it is relatively small. Not every 1940 child lived to 2000 and, of those who did, not all of them received the long form. Therefore, we are only able to match approximately 600,000 children, or 11 percent of those with PIKs. Second, those children who do match are relatively old by 2000 and past their prime working years, making it difficult to interpret socioeconomic outcomes such as labor market income and occupation measured so late in life. Third, the sample may be somewhat positively selected since individuals would have to have survived to the year 2000 and there is a well-known education gradient in mortality (e.g. Cutler, Lleras-Muney and Vogl 2011).

Therefore, to vastly increase our sample of children matched to adult outcomes, and to expand those outcomes to include income during the prime working years, we match the 1940 children to IRS 1040 income tax returns from 1974 and 1979. These files originally contained SSNs for both the main and secondary filers which enable a simple look-up in the SSA Numident file to obtain a unique PIK for both single taxpayers and members of married couples. Individuals from 1940 will not match to these data if they die before 1974 or if they did not file a tax return in either of these years. Our matched sample includes 4.63 million individuals, or 82 percent of those who were assigned PIKs. From these data we obtain information about adjusted gross income, wage and salary income, flags for having filed a Schedule C, D or SE, and place of residence.

[18] The long form was mailed to one out of six people in the U.S. in 2000.

[19] In earlier analyses, we also examined effects on home ownership, house values, disability status, and income reported in the 2000 Census (before we had access to the tax data). Given everyone in our sample was born before the 1940 Census, the 2000 Census outcomes are measured late in life, making them less useful as indicators of socioeconomic success. For example, the majority of our sample are retirees and the vast majority are homeowners by 2000.

In order to reduce yearly variability, we use the mean of inflation adjusted (in 2000$) income from 1974 and 1979 when both years of income are available.[20] We keep individuals in the sample if one year of income is missing and ignore the missing year. We also create an indicator variable for being in the tax return data in 1974/79 conditional on being PIKed in the 1940 Census to check that our sample is representative.[21]

Our migration and neighborhood characteristic outcomes are computed from the latitude and longitude of the address of the tax filer and the 1940 Census household.[22] In particular, 1940 Census and 1974 and 1979 IRS residences are mapped to a 1970 Census tract, with the sample restricted to those we can place in a 1970 Census tract. This restriction turns out not to matter for the 1940 sample, as nearly all of our HOLC cities had a large enough population that they were tracted by 1970. However, by 1974 and 1979, about 15 percent of people had moved to places that were not tracted by 1970, which reduces our sample size relative to the sample used to analyze income outcomes. We consider six Census tract-level neighborhood measures: the college graduation rate, the high school graduation rate, the employment to population ratio, the unemployment rate, the poverty rate, and the share of households with a single head. We also construct a composite neighborhood characteristic index that is aligned so that a higher value indicates a worse outcome by taking the z-score of each variable, adding the first three z-scores, subtracting the last three z-scores, and dividing by 6.[23]

Table 1 provides summary statistics for our outcomes of interest in both the tax data (Panel A) and the 2000 Census (Panel B) for the different subsamples that we will use in our estimation, which we turn to next.

---

[20] The inflation adjustment is the Consumer Price Index for all urban consumers (CPI-U). We keep any zeroes that are reported for wage and salary income and any zeroes or negative values reported for adjusted gross income.
[21] As we show later, we find that there is a very slightly lower probability of being in the tax data for children who grew up on the D-side of a D-C boundary.
[22] Geolocation of addresses in both datasets was done in the recent past (since 2015) and may be more likely to fail for street names that have changed.
[23] We subtract the last three *z*-scores because for those measures a higher value is generally viewed as an adverse outcome.

Identification Strategy

It is well established in the historical record that the HOLC defined neighborhoods based on the observable characteristics of areas and trends in these characteristics over time. Therefore, a simple comparison of the outcomes of children growing up in neighborhoods with different HOLC grades might reflect these pre-existing characteristics and trends rather than capture the causal effects of the neighborhood. Therefore, a major challenge to any analysis of the maps is developing an identification strategy that can yield plausible causal estimates.

Our empirical approach follows two of the estimation strategies used in AHM and Aaronson, Faber, Hartley, Mazumder, and Sharkey (2021). Both strategies start with the sample of 1940 Census children living in one of the 149 cities for which we have maps. We further restrict the sample to children who lived in the buffer zones (within ¼ mile) on the two sides of the D-C and C-B boundaries. Narrowing the sample to just those living within a few city blocks of a boundary, reduces, but doesn't eliminate, the stark differences in the housing and other characteristics of families by neighborhood grade. For example, these families would have shared many amenities (e.g. transportation, labor markets) in common despite living in different neighborhoods. At the same time, we know from the historical record and from statistical analysis in AHM that these neighborhood characteristics sometimes changed sharply at neighborhood borders, thereby making simple cross-border comparisons fraught with bias.

Method 1: Use of comparison boundaries

Our first strategy uses propensity score methods to create a set of "comparison" boundaries that are weighted to resemble the actual "treated" HOLC boundaries. The intuition behind this approach is that there were likely many "missing" boundaries that the HOLC could have used to separate neighborhoods but did not for practical reasons. For example, buried within a largely homogenous D neighborhood might be a small pocket resembling a C neighborhood based on demographic and housing characteristics. But it may not have made practical sense to classify this smaller area nestled inside a larger neighborhood as a separate neighborhood.[24] Our strategy uses children who grew up in the buffer zones on the two sides of these potential boundaries as a comparison group. We can then compare the cross-boundary *differences* in the treated sample of children to the cross-boundary *differences* in our comparison boundaries. This difference-in-differences strategy allows us to infer the true causal effect of growing up on the lower-graded side of a redlined or yellowlined neighborhood.

To implement this strategy in practice, we overlay a one-half mile by one-half mile grid over each city. We then take all the grid segments that are located within HOLC neighborhoods of the relevant grade. For example, for the estimation of the effects along D-C boundaries, all of the grid segments inside of D areas and C areas are used to construct a comparison sample of boundaries and buffer zones. For each comparison boundary, we randomly assign one side to be the lower-graded side (lgs).[25]

Next, we use the propensity score model below to develop weights that will be used to ensure that, pre-map, comparison group segments are similar to treated boundaries based on observable characteristics:

$$1\{Treated\}_{b,c} = \alpha_c + \sum_{k=1}^{K} \beta_{1910}^{k} z_{b,c}^{k,1910} + \beta_{1920}^{k} z_{b,c}^{k,1920} + \beta_{1930}^{k} z_{b,c}^{k,1930} + \epsilon_{b,c}.$$

Each observation is a border segment *b* located in city *c*. α_c is a city fixed effect and $z_{b,c}^{k,t} = x_{lgs,b,c}^{k,t} - x_{hgs,b,c}^{k,t}$ is the gap between variable *k* on the lower-graded side (*lgs*) versus the higher-graded side (*hgs*). The sample pools treated and comparison borders and $1\{Treated\}_{b,c}$ is an indicator equal to one if the border is treated. The variables in *k* include: share African American, share foreign born, African American population density, White population density, homeownership rate, log house value, log rent and the share of homeowner households that have a mortgage.

[24] See Appendix Figure A4 in AHM for a stylized example of this situation. Figure 2 in AHM shows that this is especially plausible for Chicago where there are many broad swaths of red-shaded neighborhoods that likely contained pockets that resembled yellow graded neighborhoods.
[25] We use this approach to make sure that the distribution of boundary differences in our comparison set of borders is representative of the underlying universe of all such boundaries and is not overrepresenting either tail of the distribution. The reweighting of comparison boundaries occurs after the randomization.

We use the results of the propensity score model to create a set of inverse probability weights (IPW) for each individual with the following specification:[26]

$$y_{igbt} = \beta_t 1[lgs]1[treated] + \beta_{lgs}1[lgs] + \beta_{treated}1[treated] + X_{igb1940} + \alpha_b + \epsilon_{igbt}.$$

In this equation, an observation is a person, i, living within ¼ miles of a border segment, b, on either the lower- or higher-graded side, g, in 1940. The sample pools individuals living in the buffers of the actual treated borders along with those living in the buffers of the comparison borders. The index t reflects the fact that the outcomes are measured in either the 1970s (IRS data) or 2000 (Census). $1[treated]$ is an indicator for a treated border, $1[lgs]$ is an indicator for the lower-graded side of the border, α_b is a border segment fixed effect, and $X_{igb1940}$ is a set of covariates from the 1940 Census which were likely determined prior to the HOLC maps. They include the race of the child, whether the child was a teen birth, family size indicators, and mother and father measures for age, marital status, race, foreign born, citizen, Hispanic, and educational attainment.

AHM find strong evidence that this approach largely eliminates any gaps in the cross-boundary differences between the treatment and comparison groups in the share African American, home ownership rate, house values, and rents in the pre-period before the maps were drawn which provides support for this identification strategy.[27]

[26] The weights are as follows: for the comparison boundaries, w = pscore/(1-pscore) and for the "treated" boundaries, w = 1. The sample excludes treated borders with a propensity score above that of the maximum comparison border and comparison borders with a propensity score below that of the minimum treated border. Using this procedure essentially "up-weights" comparison boundaries that are most similar to treated boundaries and "down-weights" those that are least similar. As a result, the reweighted comparison borders look more similar to the actual HOLC borders than the unweighted comparison borders do. AHM also show that there is a sizable amount of overlap in the distributions of the propensity scores for the two groups (see their Appendix Figure A6, Panels A and B).

[27] We also validated this approach using children 16 years and younger in the 1910, 1920, and 1930 Censuses. We ran regressions using our two main estimation strategies and used the following left hand side variables to check for balance in the pre-period: an indicator for being a Black person, an indicator for living in an owner-occupied household (1920 and 1930 only), the log of rent if living in a rented home, and the log of the home value if living in an owned home. Again, we were largely able to eliminate gaps in these outcomes in the period before the maps were drawn as the estimates in this pre-period were generally economically small and statistically insignificant. The two exceptions are that the estimate for the 1930 gap in rent along the C-B boundary is -3 percent and statistically significant at the 10 percent level and the estimate for the 1920 gap in owner-occupancy along the D-C boundary is -2 percentage points

Method 2: Low Propensity Score (LPS) Borders

Our second approach is based on the idea that some borders may have been drawn for more idiosyncratic reasons and did not actually reflect meaningful differences between adjacent neighborhoods of different grades. One way these "misaligned" borders could arise is because the map makers may have simply needed to close a polygon, and the exact street that was chosen to do so was somewhat arbitrary.[28] The way we operationalize this idea is to limit our sample to the subset of actual HOLC borders that appear "idiosyncratic" -- that is, where meaningful pre-existing differences across boundaries do not exist. A nice advantage of this approach is that, by definition, it no longer requires an explicit comparison group based on counterfactual boundaries, and therefore provides a second approach that complements our first strategy.

Practically, we use the predicted p-scores from our model and select the actual HOLC borders with a low predicted probability of being chosen (below median p-score) based on our model.[29] AHM show that this approach effectively removes pre-existing cross-boundary differences and trends providing supportive evidence for the validity of this method. The specification for this model is:

$$y_{igbt} = \beta_t 1[lgs] + X_{igb1940} + \alpha_b + \epsilon_{igbt}.$$

4. Results

Education

We begin by describing our findings on educational attainment based on the link between the 1940 and 2000 Census. Figure 1 shows four estimates of the effects of living on the lower-graded side of an HOLC boundary on completed years of schooling, corresponding to the two border types (D-C and C-B) and the two estimation strategies. "T vs C" is shorthand for treated borders vs comparison borders (method 1) and "LPS" refers to low propensity score boundaries (method 2). One striking finding is the consistency

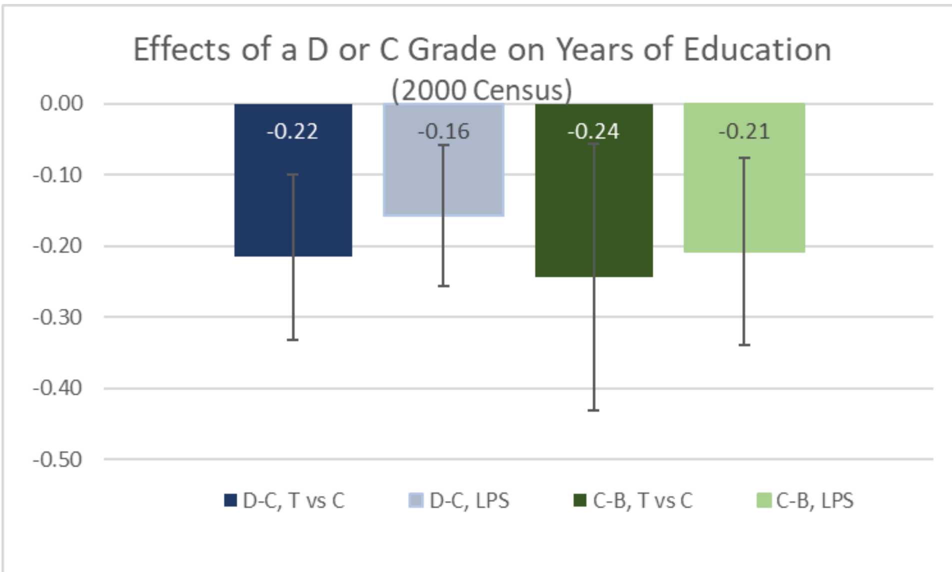
and statistically significant at the 5 percent level. However, we did not correct for multiple hypothesis testing, and since we ran a total of 16 regressions the exceptions could be due to chance.

[28] See Appendix figure A4 and the text within AHM for a stylized and actual example of this situation.

[29] Our focus on low-p-score boundaries is akin to the subclassification approach described in Imbens (2015) and Imbens and Rubin (2015).

of the estimates; growing up on the lower-graded side of a neighborhood leads to an approximately 0.2 decrease in years of education for both D-C and C-B comparisons using either estimation strategy. These effects are in all cases highly statistically significant, suggesting that something about the lower neighborhood grade led to worse educational outcomes. Relative to a mean of roughly 13-14 years, 0.2 years translates to a 1.5 percent reduction in years of schooling.

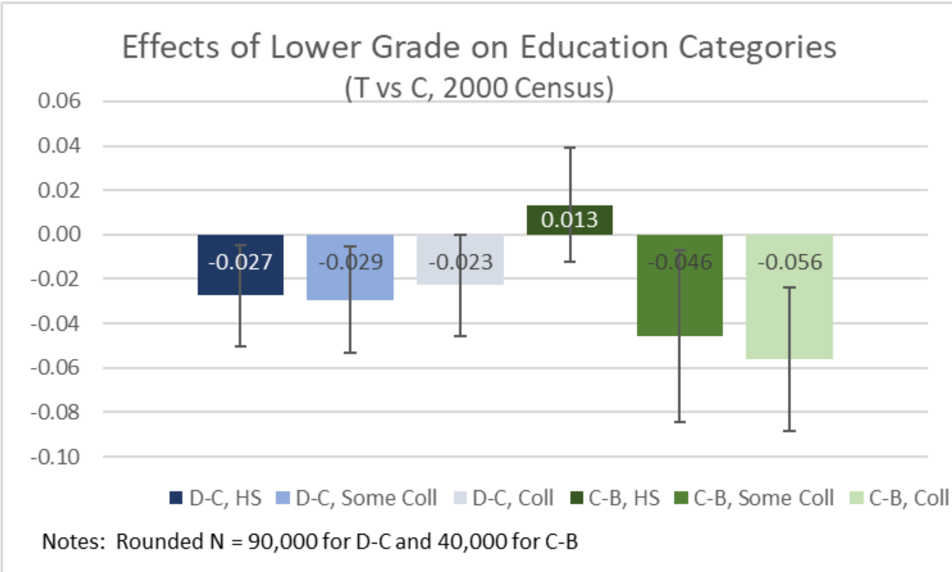
Figure 1



We also examined whether the education effects occurred at a particular stage of schooling (HS, some college, or college). The results for our treatment versus comparison borders are shown in Figure 2.[30] Along the D-C borders (blue bars), we find a consistent 2 to 3 percentage point reduction across all 3 education margins. However, along the C-B borders (green bars), we find no effect on the high school margin but a much larger 4½ and 5½ percentage point reduction in attending college and graduating college, respectively. These reductions are economically significant compared to the mean college attendance rate of 53 percent and the mean college graduation rate of 26 percent for this sample.

[30] The estimates using the LPS method are generally smaller for high school and some college for both border types but are similar for college or more.

Figure 2

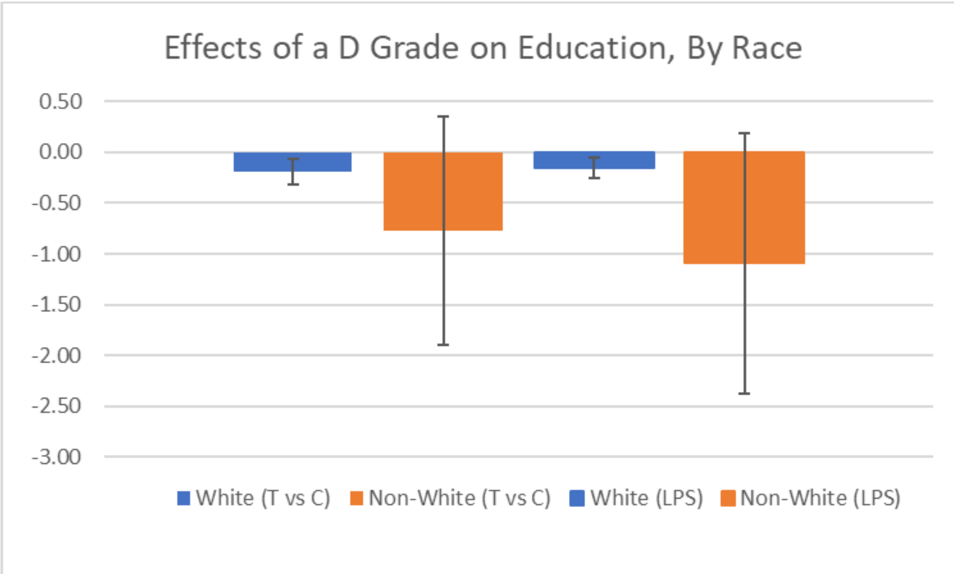


We also estimated separate effects for Whites and non-Whites and men and women.[31] We found much larger negative point estimates for non-White children, as shown in Figure 3, but these estimates came with extremely large standard errors, leading us to conclude that our samples were too small to allow for sufficient precision to detect significant racial differences.[32]

Figure 3

[31] We grouped all non-White peoples together for the purposes of meeting Census Bureau disclosure avoidance review requirements.

[32] Putting aside precision, the differences in point estimates by gender did not follow a consistent pattern along D-C boundaries but men were more negatively affected along C-B boundaries. However, none of these gender differences were statistically significant.



Income

Next, we turn to results on mid-career income reported on tax forms, starting with wage and salary income in Figure 4. We find large, negative, and statistically significant effects across three of our four estimates. Unlike with years of schooling, the magnitude of the results appears to differ by border type and by the choice of estimation method. Along the D-C borders, estimates range from a loss of $1,120 when using only low propensity score borders to a loss of $2,140 when comparing the treated borders to comparison boundaries. The analogous estimates along the C-B boundaries are a loss of $862 and a loss of $2,760. Relative to the mean level of wage and salary income in these samples, the effects range from -1.3 percent to -4.6 percent.

Figure 4:

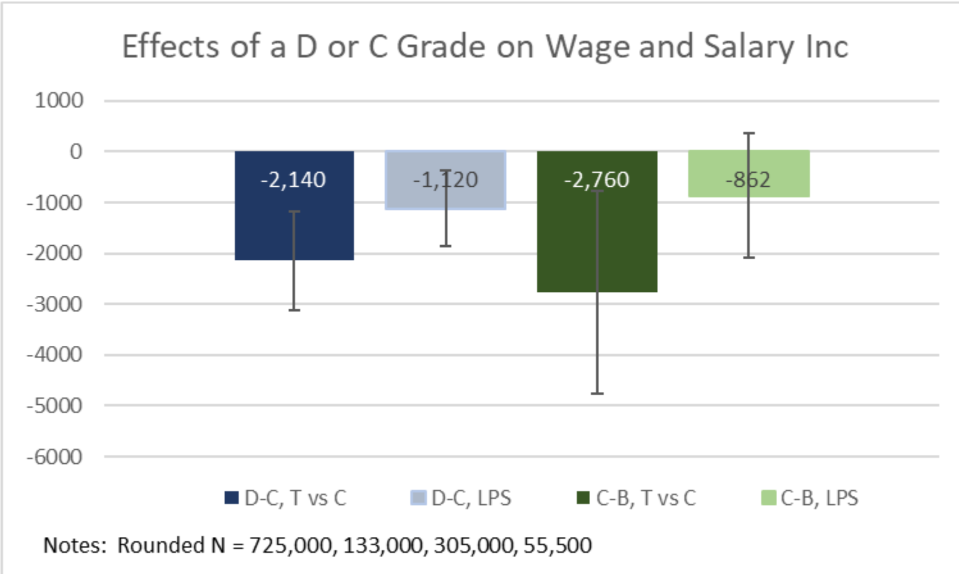
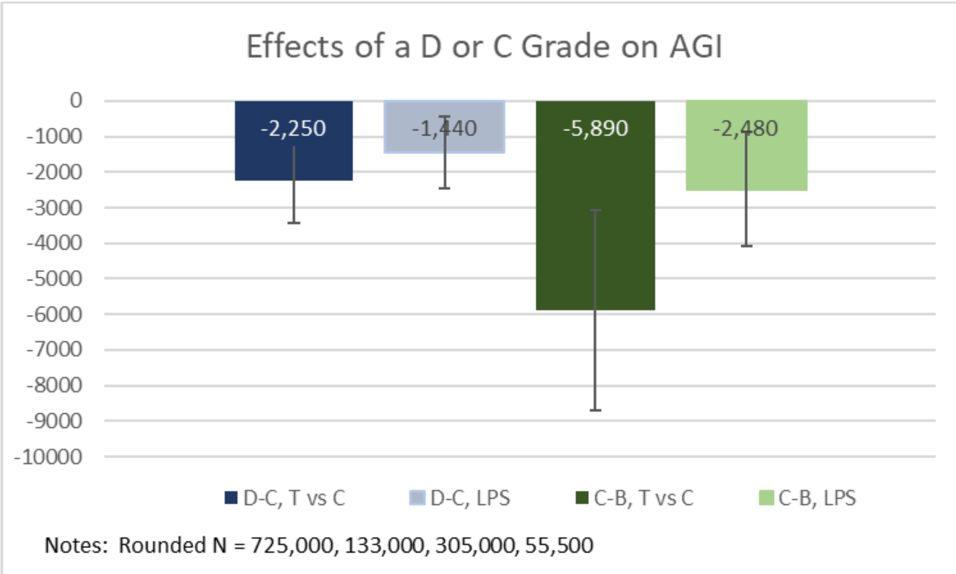


Figure 5 displays estimates on Adjusted Gross Income (AGI), which allows for other forms of income in addition to wages and salaries. Here we find even larger negative effects that are highly statistically significant. Along the D-C border, our estimates range from -\$1,440 to -\$2,250, which are modestly higher than what we found using just wage and salary income. However, our AGI estimates along the C-B boundaries are at least twice as large (-\$2,480 and -\$5,890) as comparable estimates using just wage and salary income. Larger estimates along the C-B boundary are consistent with neighborhood level housing outcomes in AHM and suggest that the effects of yellowlining may have been especially severe for some. We return to these estimates later in the section on heterogeneity.

Figure 5:



Migration and neighborhood characteristics in adulthood

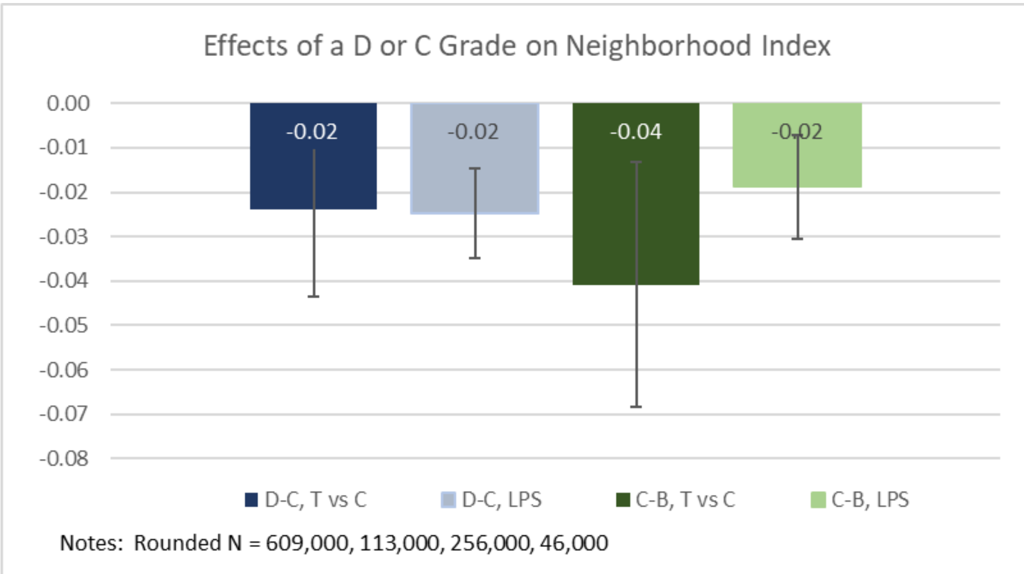
In this section, we consider both the migration response to HOLC grades as well as measures of the neighborhood where children were living later in life. Roughly 96 to 97 percent of our sample of children in 1940 moved to a new Census tract by 1970. However, only 50 to 60 percent moved to a new county and 25 to 30 percent to a new state. We ran our main specifications using these three migration measures as outcomes and found that both redlining and yellowlining led to no effect on changing tracts (since nearly everyone had moved tracts) and only modestly lower state and county migration rates of about 0.5 and 2 percentage points.[33]

However, there was a statistically robust and economically relevant difference in the characteristics of the places where they moved. Figure 6 shows the effect of redlining and yellowlining on the composite neighborhood index of the location where the children were living in adulthood during the 1970s. Effects range from -0.02 to -0.04 standard deviations, depending on the specification. By way of one comparison, Kling et al (2007) find that the effects of the Moving to Opportunity program increased their summary

[33] The reductions in migration are not statistically significant for the D-C border when using our treatment versus comparison group specification, but they are statistically significant for the other three specifications.

measure of the socioeconomic outcomes of adults by 0.04 standard deviations and that of youth by 0.02 standard deviations.[34]

Figure 6:



In figure 7, we show each of the components of the neighborhood index using the treatment versus comparison group method and the redlining sample.[35] It is immediately evident that the educational components of the index are driving the effect on the composite neighborhood index. Children that grew up on the D side of a D-C border in 1940 live in Census tracts in the 1970s that have 0.05 standard deviations lower rates of attainment of college or more and high school or more. We also find economically smaller effects on the neighborhood poverty rate, the unemployment rate, and the rate of single headed households of around 0.01 to 0.02 standard deviations. Although these estimates are not statistically significant, they are significant, and of a similar magnitude, when using the low propensity score method (not shown).

Figure 8 reveals a similar pattern for the effect of yellowlining on the components of the neighborhood composition index. Again, lower educational attainment is driving the drop in the composite

[34] This is from Kling et al (2007)'s intent to treat estimates of the experimental group versus the control group (E-C) on summary measures shown in their Table 2.

[35] Each outcome is measured in standard deviation units and aligned so that a negative sign indicates an adverse outcome.

neighborhood index. In this case, the effects are even larger, with an estimated decline of around 0.07 to 0.09 standard deviations in our educational outcomes and an increase in rates of neighborhood poverty, unemployment, and single headed households of around 0.03 to 0.05 standard deviations.

Figure 7:

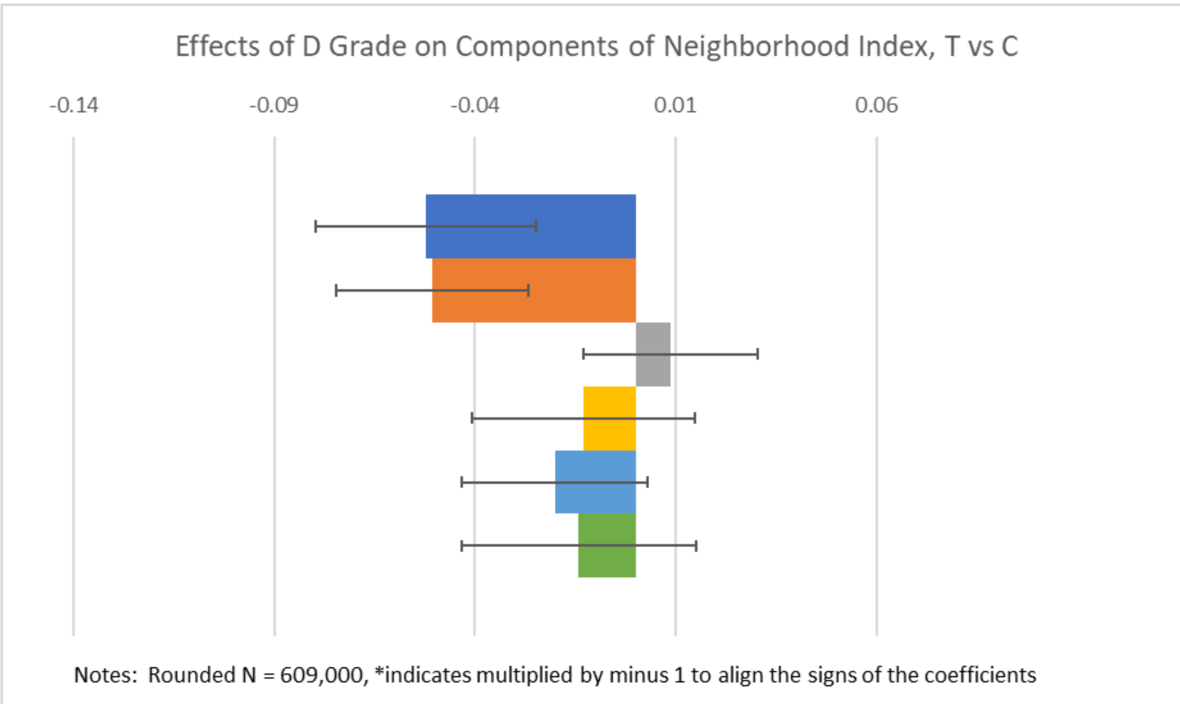
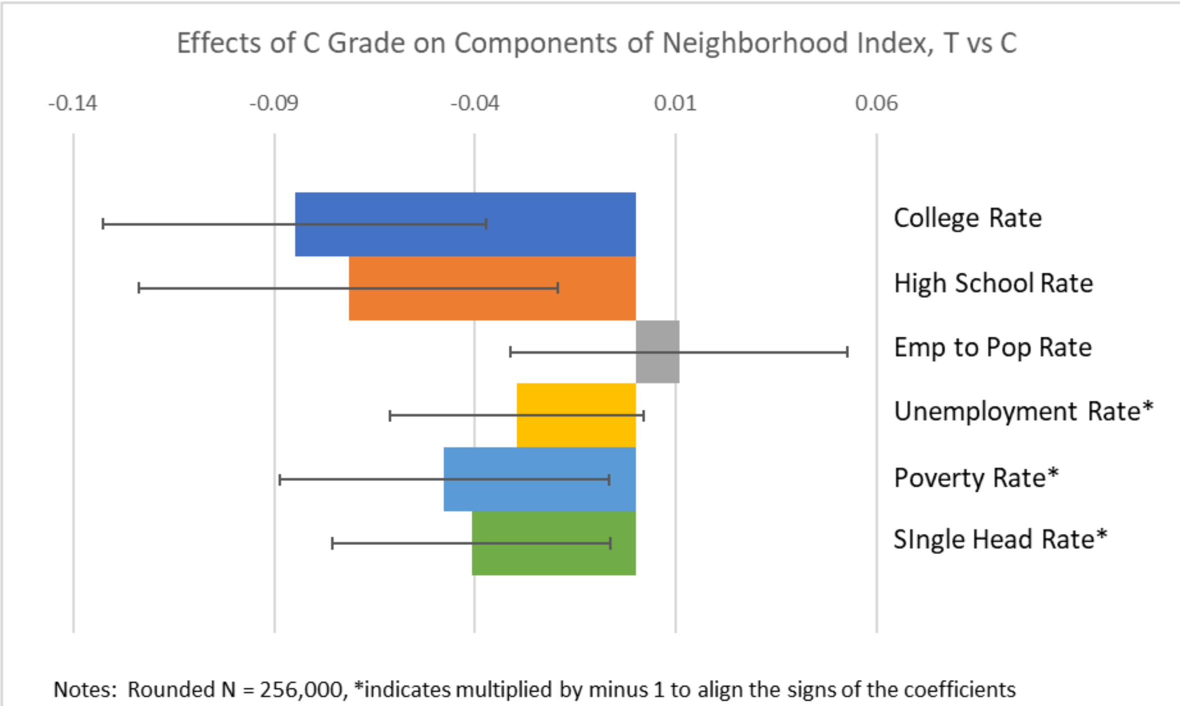


Figure 8:



Heterogeneity of Effects on AGI

Next, we stratify the sample by demographic and socioeconomic characteristics to consider situations that may have led to especially negative outcomes and to begin to collect results that could help uncover mechanisms. To help with precision, we focus on the tax records where our sample is substantially larger. Even still, sample size ultimately limits how much can be inferred about many group differences. We also focus on AGI as our outcome of interest, partly for brevity, but also to highlight some useful differences by the type of income filers.

We begin by looking at individuals who owned a business or had financial income. Specifically, we use an indicator variable for whether individuals filed either a schedule C or schedule SE to proxy for whether they had business income. Similarly, we use an indicator for whether individuals filed a schedule D to capture capital income from sources such as financial or real estate investments. We also construct a separate indicator for those who fit into both categories. Finally, we compare these three groups to those who did not file C, D, or SE schedules.

The results using the treated vs comparison group approach are shown in Figures 9 and 10.[36] We find an especially large negative effect on AGI among those with schedule D capital income. Along D-C boundaries, the magnitude of the effect is -$3,400 for those who only filed a schedule D for capital income, although there is no significant effect for those who file both a business schedule and schedule D. The impact is extremely large along the C-B boundaries: -$10,100 for those who only file a schedule D and -$14,700 for those who filed both a business schedule and a capital income schedule. These results are consistent with individuals experiencing smaller capital income gains from housing investments if they grew up in lower-graded neighborhoods.

Figure 9

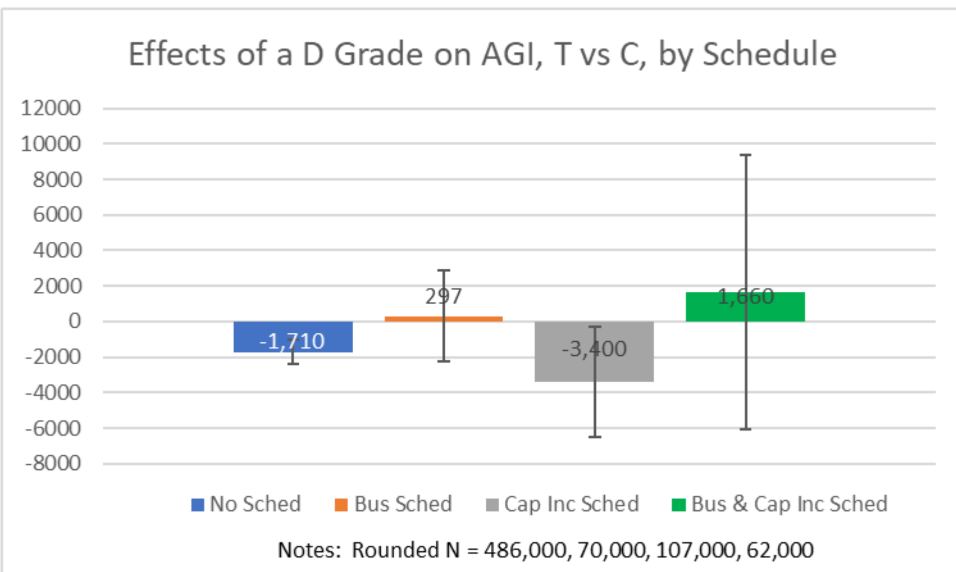
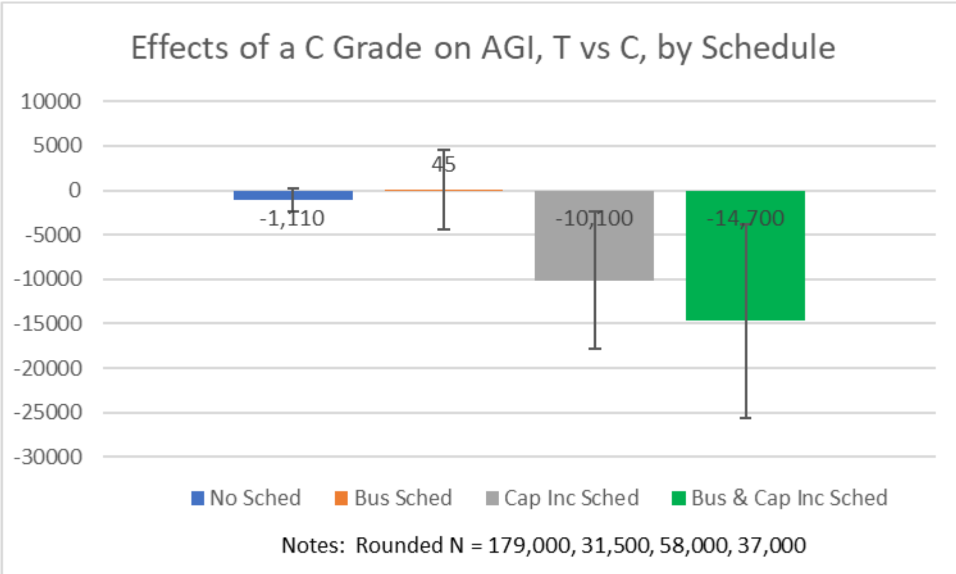


Figure 10:

[36] We find broadly the same qualitative patterns when using the low propensity score method. However, the magnitudes of the estimates differ in a similar pattern to the full sample shown in Figure 5.



We now turn to examining heterogeneity based on characteristics from the 1940 Census. Figures 11 and 12 provide treated vs comparison group estimates by homeownership, age, income of parents in 1940, and whether the parents were married and present in the household.[37] The top bar reproduces the overall effect on AGI for each border type that was previously displayed in Figure 5. It is important to note that the scales on Figures 11 and 12 are quite different, and the overall negative effects of yellowlining in Figure 11 are more than twice as large as those of redlining in Figure 12.

The first set of results considers whether the parent was a homeowner or renter in 1940. One hypothesis is that children of homeowners might have been more severely impacted by financial

[37] See our earlier discussion of differences by race and gender. We also considered differences by proximity to the border and by whether families had migrated between 1935 and 1940, but it was *a priori* unclear what such differences may reveal. Regarding proximity to the border, focusing only on those families living extremely close to the border (within 1/8th of a mile), on either side, might compare more similar families and thus might provide a better causal estimate. At the same time, there could be negative spillover effects for those on the higher-graded side if they lived extremely close to the lower-graded side. Our estimates were consistently higher for those living farther from the border (between 1/8th and ¼ mile) but the magnitudes varied quite a bit depending on our estimation strategy and border type. Regarding movers in the 1930s, those who lived at the same address in 1935 may have been "treated" for longer (depending on exactly when the maps were made in their city) and could have experienced larger effects. On the other hand, families who decided to move from 1935 to 1940 into an area that received a lower HOLC grade may reflect a selected sample. We generally found more negative income effects for children in households that had been in the same address since at least 1935 compared to those whose households had moved between 1935 and 1940. Appendix Figures A1 and A2 show comparable results using the low propensity score approach. These results generally reveal similar patterns.

disinvestment in a neighborhood due to declining access to credit. We discuss this point in greater detail in the next section. AHM show evidence that the lower-graded side of neighborhoods experienced declines in house values and suffered impacts associated with financial disinvestment such as higher rates of vacancies and dilapidated housing. Indeed, we find that the negative effects on AGI appear to have been larger, albeit not precisely so, for children whose parents were homeowners. For example, along C-B borders, the effect on AGI for children of homeowners was -$7,430 versus -$3,680 for children whose parents did not own their home. However, these differences are not statistically significant at conventional levels.

A growing literature on the effects of early life events has highlighted how the consequences of shocks experienced by young children may be especially severe (e.g. Almond and Currie, 2011, Almond, Currie and Duque, 2017). In the context of neighborhoods, Chetty, Hendren and Katz (2016) found that the Moving to Opportunity program only had effects on the long-run income of children below the age of 13. Therefore, we divided our sample into those who were "young" in 1940 (below the age of 9) and those who were "old" (between the ages of 9 and 16). We found little difference in the estimates between these age groups though the point estimates were consistently larger for the older group particularly when using the low propensity score approach (Figures A1 and A2).

We also consider how the effects might differ based on two measures of the socioeconomic status of the family in 1940, namely the income level of the parents and whether there were two married parents in the household. We created three categories of wage and salary income, those with below median income ("low"), those with above median income ("high") and those with zero or no reported income ("missing"). We find a similar pattern by parental income along both the redlining and yellowlining borders. The effects are least negative for the children with high-income parents, then more negative for the children with low-income parents, and the most negative for children whose parental income is missing in 1940. Unfortunately, the latter mixes business owners and farmers, since their income was not collected in 1940, with those who did not report an income. But there may have been reason that each group would have been hard hit by less access to credit. The maps may have been particularly detrimental to very low-income households that had no earnings. Moreover, it is also possible that the maps negatively impacted access to credit for businesses or further exposed business owners to negative real estate shocks if they owned their premises.

Finally, we see almost no difference in the effect of redlining between children with two parents in the household and those without two parents. The effect of yellowlining is more negative for children with two parents, but the sample size is somewhat small (26,500) for those without two parents, resulting in a wide confidence interval.[38]

Figure 11

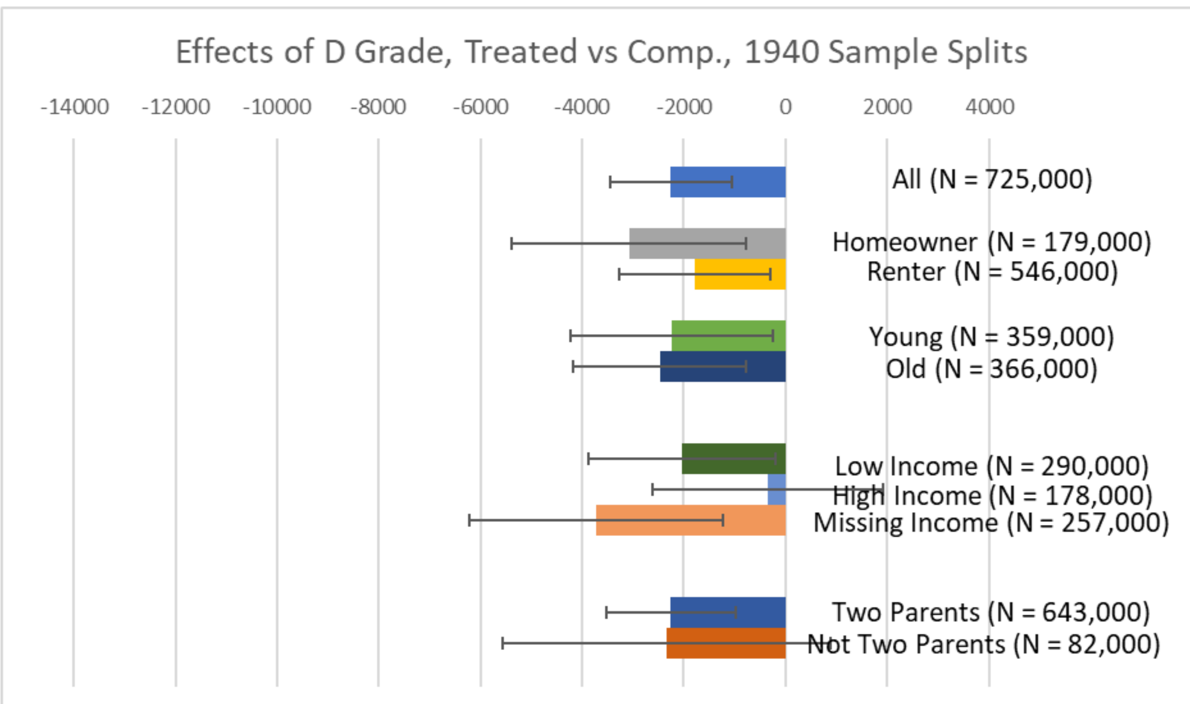
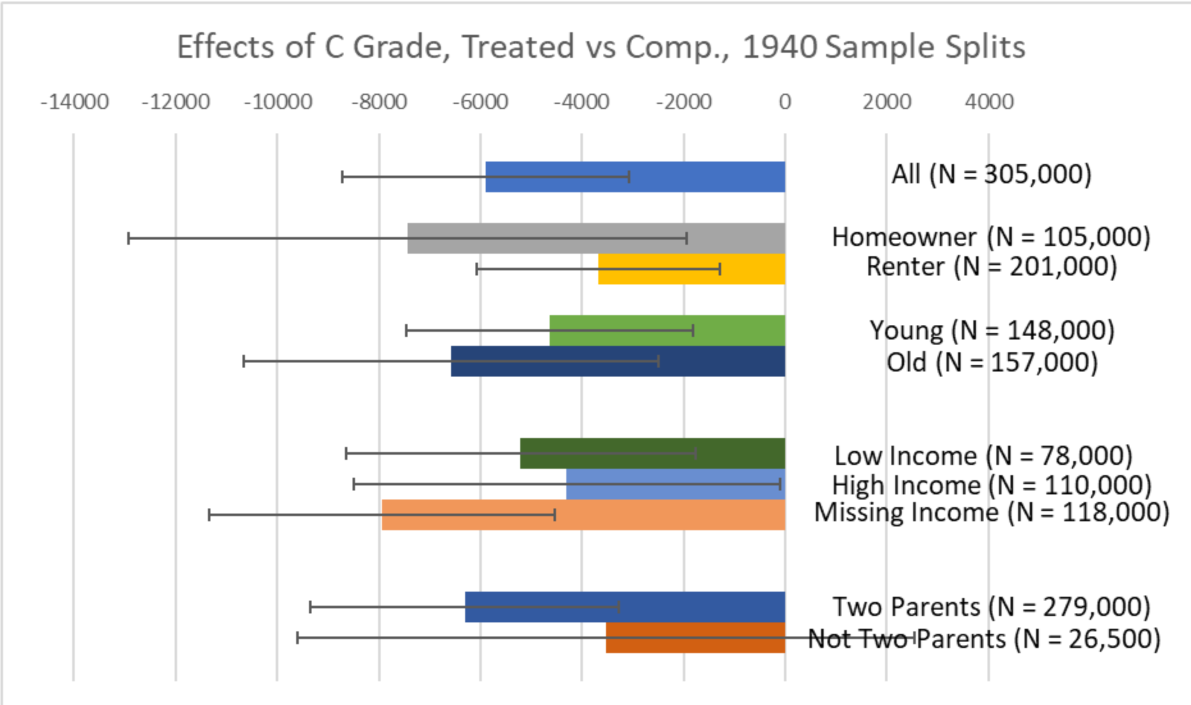


Figure 12

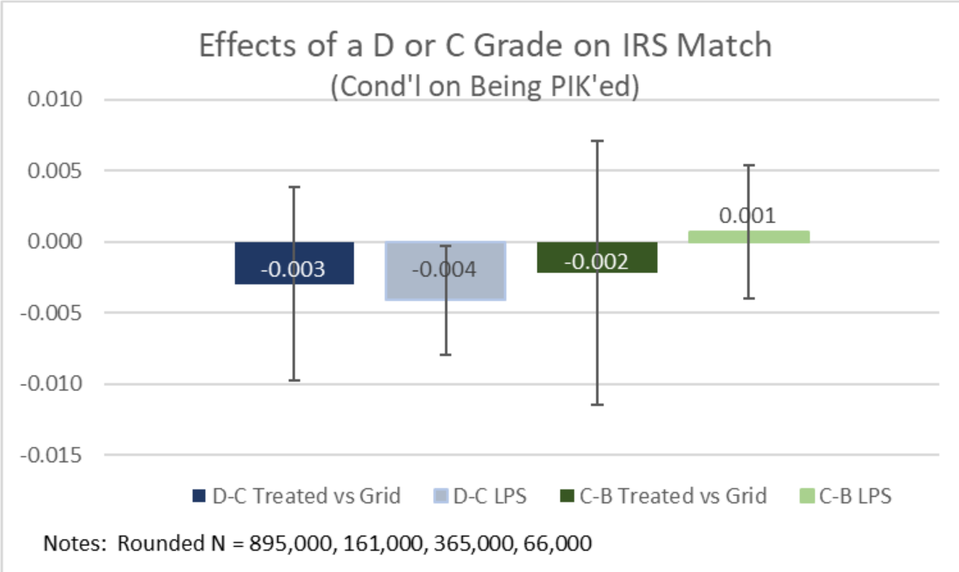
[38] We also have a relatively small subsample of 82,000 children who did not live with both parents when estimating this along the D-C boundary compared to 643,000 children who lived with both parents.



Selection into the Tax Sample

One concern is not all children in the 1940 Census who were successfully matched to the Social Security NUMIDENT database and received a personal identification key (PIK) are observed in the tax data. This selection issue could arise from children not filing taxes as adults or from any failure to match the tax records to Social Security records. Across our four estimation samples the baseline rate of being included in the tax records is between 82 and 84 percent, suggesting that we are capturing most of the potential sample. Still, in Figure 13, we show the results of using our statistical model where the outcome is an indicator of inclusion in the tax records and our universe is the sample of all children who were "PIKed." In all four models, we find very small negative point estimates of less than 0.5 percentage points and in only one case is the estimate statistically significant. This suggests that there is no meaningful difference in the effect of being redlined or yellowlined on selection into the tax sample.

Figure 13



5. Discussion

Mechanisms

We find economically significant, negative long-term consequences for children who grew up in lower-graded neighborhoods. Naturally, the critical question is why. A key starting point is that these lower-graded neighborhoods likely experienced reduced access to credit and that this led to a sharp and rapid decline in housing demand as marginal or credit constrained borrowers could no longer afford to buy a home in these neighborhoods.[39] This is supported by evidence from AHM who find that by 1940 the HOLC maps had caused home values to fall by 9 percent in "redlined" neighborhoods and 10 percent in "yellowlined" neighborhoods. They also find that the effects of the maps were to reduce rents by 7 percent along D-C boundaries and 2 percent along C-B boundaries.[40] Since the maps were largely completed by 1937, the housing supply would have had very little time to adjust by 1940. Therefore, the decline in rents and house values in these areas are likely due to a drop in housing demand that was caused by reduced access to credit. Housing demand could have also fallen due to a decline in the quality of housing units in

[39] In a more recent episode, Jensen and Johannesen (2017) find that contractions of credit supply after the financial crisis caused affected Danish households to decrease consumption, including real estate consumption.
[40] We find very similar results when we estimate these effects using individual-level data for the sample of children 16 years old or younger in 1940.

these neighborhoods as landlords began to reduce investment as the value of their properties declined due to lower HOLC grades. This could have also been accompanied by reductions in local public investment in public goods such as schools and public safety.

One potential mechanism for the decline in long-run outcomes for children living in these neighborhoods is that the decline in the value of homes and the accompanying loss of wealth, may have made it harder for parents to invest in their children's education. This may have been a factor, for example, in driving the roughly 5 percentage point reduction in college attendance for children growing up in "yellowlined" households. The reduction in children's education could then have resulted in decreased income later in life and a greater likelihood of living in neighborhoods with lower educations levels and higher poverty rates. However, if all the effects on earnings that we estimate were due solely to reductions in educational attainment, then it would imply very high rates of return to education on the order of 14 percent per year.[41] Such a rate of return might be implausibly high, or that there were other effects of reduced credit access that operated outside of the channel of human capital investment.[42]

Another possible mechanism that can explain our findings is related to the effects that the HOLC maps had on the rates of homeownership among the parents of the affected children growing up in lower-graded neighborhoods. For example, AHM estimate that redlining caused homeownership rates to fall by 1.2 percent by 1940 and 2.6 percent by 1950 and found even larger declines along C-B boundaries. This period covers the years in which our sample of children would have been living at home with their parents. Parental homeownership may provide the needed stability for later-life success for children through having positive peer influences and providing networks that could enable future opportunities or simply from better parent labor market outcomes (Hausman, Ramot-Nyska, and Zussman 2021). By contrast, disruptions in

[41]For example, our estimate of the effects of redlining (D-C borders) based on the treatment vs comparison group method imply a reduction of 0.215 years of schooling and a decline of 3.78 percent (relative to the mean) in wage and salary income. This translates into an estimate of a 17.6 percent rate of return (3.78/0.215) per year of education. A simple average of this return to education across our four estimation models is 13.7 percent.

[42] For example, Goldin and Katz (1999) using log weekly wage differentials estimate that the returns to a year of high school in 1979 for young men was 8.1 percent and the return to a year of college for young men was 8.4 percent. The corresponding rates in 1969 were 7.4 percent and 9.6 percent.

the residential, and possibly schooling, environment experienced by children could lead to instability and worse socioeconomic outcomes. That said, studies on the effects of homeownership and residential stability on children's outcomes are typically associative (e.g. Adam 2004; Green and White 1997).

Still, these disruptive housing effects are consistent with the more negative impacts on children's income for families whose parents were initially homeowners. Moreover, higher residential instability arising from redlining and yellowlining could have affected a number of outcomes besides educational attainment and therefore provide a potential explanation for why the long-run effects on income are greater than what would have been predicted based on the estimated effects on education alone. It would be useful for future work to also explore whether children's health was negatively impacted by redlining and yellowlining as this could provide another explanation for the very large, implied returns to schooling. To date, work on the health effects of the maps is purely associative and study later generations.

We also speculate that as credit access declined in lower-graded neighborhoods, there could have been disruptive effects arising from losing local businesses that may have impacted children's long-run outcomes. This may have particularly been the case for family run businesses as it was more common in 1940 for families to live above their shop. More generally, commercial real estate could have also been impacted by the sharp decline in real estate values if local business owners also owned residential properties. This would be consistent with our finding that children whose parents had missing income, some of which is due to their owning a business and therefore not having their income collected in the 1940 Census, experienced larger income losses in adulthood.

Caveats

One important downside of our focus only on children living in buffer zones around the borders of HOLC neighborhoods is that our estimates may not generalize to the children who grew up in the interior of neighborhoods. For example, it could be the case that areas in the interior of neighborhoods had higher concentrations of Black families and Black-owned businesses, and that the children growing up in these areas were much more severely impacted than the children growing up near a border and therefore in closer proximity to a higher-graded neighborhood.[43] We believe that this is an important area for future research to consider, although it will be challenging to come up with a research design that can deliver causal estimates. Similarly, there may be other important outcomes that our analysis did not consider due to data limitations such as effects on business ownership, social networks, criminal activity, health, and exposure to pollution.

It is also important to recognize that there were many other overlapping policies and environmental influences that were affecting these neighborhoods both at the time the maps were drawn, as well as in subsequent years. For example, urban renewal policies that started in the 1950s that aimed to clear "slums" and redevelop urban areas and which were sometimes associated with highway construction, may have also impacted the opportunities faced by some of the children in our sample.[44] The boundaries of school districts and school funding could have also been affected by the maps.[45] Our estimates should be viewed as the reduced form effects of all of the possible effects that were associated with redlining and yellowlining.

Finally, given that much remains unknown about the extent to which the information in the HOLC maps were used by private actors or the FHA in their decisions to insure mortgage loans (see section 2), it is important to carefully consider how to interpret our findings. As we discuss earlier, it may be useful to think of our analysis as serving as a proxy for the overall effects of discriminatory policies that targeted urban neighborhoods. If so, our estimates may serve as a *lower bound* of the full effects of such policies. This is because the grade classifications and the borders between neighborhoods chosen by the HOLC maps are unlikely to have always lined up with those chosen by other pertinent actors, adding measurement error and likely attenuating estimates.

6. **Conclusion**

[43] We thank Trevon Logan and Sun Kyoung Lee for their comments in this spirit.
[44] See Collins and Shester (2013), LaVoice (2019) and Shi et al (2021) for recent studies of urban renewal.
[45] See Lucas and Cleveland (2021).

The federal government's initial attempt to institute a national housing appraisal system in the mid to late 1930s was meant to help public and private lenders better understand community-level repayment risk, a worthwhile goal following the devastating foreclosure crisis earlier in the decade. However, the new system used some measures that were not tied directly to creditworthiness, including ethnicity, immigration status, and race, causing harm to poorer and more diverse neighborhoods. Indeed, past research, by us and others, has documented the geographically inequitable development that arose over the following decades.

This paper adds to this literature by studying the later life educational attainment and income of individuals who grew up in differentially graded neighborhoods when the HOLC maps were first drawn, regardless of where they moved later in life. Our approach therefore more closely resembles the modern literature that uses experimental or quasi-experimental methods to trace the impact of specific policies that change neighborhood exposure during childhood. We find economically and statistically significant effects of being on the lower-graded side of either a D-C or C-B boundary, including 0.2 years less of schooling, $1,500 to $6,000 (in 2000$) lower prime-age annual income, and neighborhood residences during adulthood with lower education and other socioeconomic characteristics. The effects on income appear to be especially large along C-B boundaries, among individuals with capital income (schedule D filers), and among children whose parents were homeowners in 1940. In light of these results, we speculate on mechanisms that may have caused access to credit of parents to have had a lifetime impact on their children. We want to stress that there is much left to be learned, however. For example, we know very little about how a variety of mid-century housing and urban policies, including the broader use of private and public appraisal maps, intertwined to have long-run implications on poorer and diverse populations. Therefore, we strongly believe this is a rich and important area of future study.

References

Aaronson, Daniel, Daniel Hartley and Bhashkar Mazumder, 2021, “The Effects of the 1930s HOLC “Redlining” Maps,” *American Economic Journal: Economic Policy*, 13 (4): 355-392.

Aaronson, Daniel, Jacob Faber, Daniel Hartley, Bhashkar Mazumder and Patrick Sharkey, 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

Adam, Emma K. 2004. “Beyond Quality Parental and Residential Stability and Children’s Adjustment,” *Current Directions in Psychological Science* 13 (5): 210–213.

Alexander, J. Trent, Todd Gardner, Catherine G. Massey and Amy O’Hara, 2015, “Creating a Longitudinal Data Infrastructure at the Census Bureau,” Working paper, U.S. Census Bureau.

Almond, D. and Currie, J. (2011) Human Capital before Age Five. In the Handbook of Labor Economics, edited by Ashenfelter, O. and Card, D. Vol. 4 (Chapter 15): 1315-1486.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature*, 56 (4): 1360-1446.

Anders, J. (2021). The Long Run Effects of De Jure Discrimination in the Credit Market: How Redlining Increased Crime. Working Paper.

Appel, Ian and Jordan Nickerson, 2016, “Pockets of Poverty: The Long-Term Effects of Redlining,” Available at SSRN: https://ssrn.com/abstract=2852856 or http://dx.doi.org/10.2139/ssrn.2852856.

Chetty, R., Friedman, J.N., Hendren, N., Jones, M.R. and Porter, S.R., 2018. The opportunity atlas: Mapping the childhood roots of social mobility (No. w25147). National Bureau of Economic Research.

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

Chyn, Eric, 2018, “Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children," *American Economic Review*, 108 (10), 3028-3056.

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

Collins, William J. and Katharine L Shester. 2013. Slum Clearance and Urban Renewal in the United States. *American Economic Journal: Applied Economics*, 5(1):239–273.

Cutler, D.M., Lleras-Muney, A. and Vogl, T., 2011, “Socioeconomic Status and Health: Dimensions and Mechanisms,” in S. Glied and P.C. Smith (eds.), Oxford Handbook of Health Economics, Oxford University Press, Oxford.

Fishback, Price V., Jonathan Rose, Kenneth A. Snowden and Thomas Storrs, 2021, “New Evidence on Redlining by Federal Housing Programs in the 1930s.” NBER Working paper 29244.

Goldin, Claudia and Lawrence Katz, 1999, “The Returns to Skill in the United States Across the 20th Century”. NBER working paper 7126

Green, Richard K., and Michelle J. White. 1997. “Measuring the Benefits of Homeowning: Effects on Children,” *Journal of Urban Economics* 41 (3): 441–461.

Greer, James, 2012, “Race and Mortgage Redlining in the U.S.,” Working Paper.

Greer, James. (2014) “Historic Home Mortgage Redlining in Chicago.” *Journal of the Illinois State Historical Society*, 107(2), pp. 204-233.

Hausman, Naomi, Tamar Ramot-Nyska, and Noam Zussman, 2020, “Homeownership, Labor Supply, and Neighborhood Quality,” Working paper.

Hillier, Amy, 2005, “Residential Security Maps and Neighborhood Appraisals: The Home Owners’ Loan Corporation and the Case of Philadelphia,” *Social Science History* 29(2), 207-233.

Hillier, Amy, 2003, “Redlining and the Home Owners’ Loan Corporation,” *Journal of Urban History* 29(4), 394-420.

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.

Hynsjö Disa M. and Luca Perdoni, 2022, “The Effects of Federal “Redlining” Maps: A Novel Estimation Strategy” working paper, Yale University.

Imbens, Guido, 2015, “Matching Methods in Practice: Three Examples,” *Journal of Human Resources* 50, 373-419.

Imbens, Guido and Donald Rubin, 2015, Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction, Cambridge University Press.

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

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.

Jensen, Thais Lærkholm, and Niels Johannesen. 2017. "The Consumption Effects of the 2007–2008 Financial Crisis: Evidence from Households in Denmark." *American Economic Review*, 107 (11): 3386-3414.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. “Experimental Analysis of Neighborhood Effects.” *Econometrica* 75 (1): 83–119

Krimmel, Jacob, 2017, “Persistence of Prejudice: Estimating the Long Term Effects of Redlining,” Working Paper, University of Pennsylvania.

LaVoice, Jessica. 2019. The Long-Run Implications of Slum Clearance: A Neighborhood Analysis. Working paper.

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

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

Massey, Catherine G., Katie R. Genadek, J. Trent Alexander, Todd K. Gardner, Amy O’Hara (2018). “Linking the 1940 Census with modern data,” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 51:4, 246-257 https://doi.org/10.1080/01615440.2018.1507772

Michney, Todd M., 2021, “How the City Survey’s Redlining Maps Were Made: A Closer Look at HOLC’s Mortgagee Rehabilitation Division” Journal of Planning History, 1-29.

Minnesota Population Center and Ancestry.com. IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]. University of Minnesota, 2013.

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.

Sagalyn, Lynne Beyer. (1980) "Housing on the Installment Plan: an Economic and Institutional Analysis of Contract Buying in Chicago." PhD Dissertation, University of Chicago.

Shi, Ying, Hartley, D., Mazumder, B. and Rajan, A. (2021). The Effects of the Great Migration on Urban Renewal. Federal Reserve Bank of Chicago Working Paper 2021-4.

Wagner, D. and M. Layne (2014). "The Person Identification Validation System: Applying the Center for Administrative Records and Research and Applications' Record Linkage Software," Center for Administrative Records Research and Applications Report Series (#2014-01).

Woods, Louis Lee, 2012, "The Federal Home Loan Bank Board, Redlining, and the National Proliferation of Racial Lending Discrimination," *Journal of Urban History* 38(6), 1036-1059.

Xu, Wenfei. (2021). "Legacies of Institutionalized Redlining: A Comparison Between Speculative and Implemented Mortgage Risk Maps in Chicago, Illinois." *Housing Policy Debate*, 1-26

Table 1: Summary Statistics

Panel A. Income Tax Data from 1974/1979

	Treated vs Comparison						Low Propensity Score					
	D-C			C-B			D-C			C-B		
	Mean	S.D.	Rounded N	Mean	S.D.	Rounded N	Mean	S.D.	Rounded N	Mean	S.D.	Rounded N
Adj Gross Income (74/79 avg)	64,110	67,340	725,000	68,940	61,480	305,000	66,050	76,820	133,000	78,600	75,020	55,500
Wage and Salary Inc (74/79 avg)	56,610	47,640	725,000	60,040	43,930	305,000	58,630	44,430	133,000	66,440	55,530	55,500
AGI, No Schedule	54,640	31,660	486,000	57,890	29,440	179,000	56,630	28,560	87,500	60,860	32,960	30,500
AGI, Bus. Schedule (C/SE)	53,100	32,230	70,000	54,910	37,640	31,500	54,750	31,900	13,000	59,260	35,510	5,800
AGI, Capital Inc Schedule (D)	94,130	79,660	107,000	96,110	73,710	58,000	95,090	82,970	21,000	111,500	97,700	11,500
AGI, both Bus. and Cap Inc Sched	94,310	164,600	62,000	103,000	130,000	37,000	96,550	208,500	12,000	115,700	130,400	7,500
Neighborhood Index (Tract)	-0.046	0.76	609,000	0.12	0.64	256,000	0.05	0.65	113,000	0.22	0.59	46,000
College Rate (Tract)	0.14	0.11	609,000	0.16	0.12	256,000	0.14	0.11	113,000	0.18	0.13	46,500
Hish School Rate (Tract)	0.61	0.17	609,000	0.65	0.17	256,000	0.61	0.17	113,000	0.67	0.15	46,500
Employment to Pop Rate (Tract)	0.58	0.06	609,000	0.59	0.06	256,000	0.59	0.06	113,000	0.59	0.06	46,500
Unemployment Rate (Tract)	0.04	0.02	609,000	0.03	0.02	256,000	0.03	0.02	113,000	0.03	0.02	46,500
Poverty Rate (Tract)	0.06	0.06	609,000	0.05	0.04	256,000	0.05	0.04	113,000	0.04	0.03	46,000
Single Headed HH Rate (Tract)	0.11	0.08	609,000	0.10	0.06	256,000	0.10	0.07	113,000	0.09	0.06	46,000
Moved Tract	0.96	0.19	609,000	0.96	0.19	256,000	0.97	0.18	113,000	0.97	0.18	46,500
Moved County	0.51	0.50	609,000	0.54	0.50	256,000	0.56	0.50	113,000	0.62	0.49	46,500
Moved State	0.28	0.45	609,000	0.30	0.46	256,000	0.26	0.44	113,000	0.34	0.47	46,500

Panel B. 2000 Census

	Mean	S.D.	Rounded N	Mean	S.D.	Rounded N	Mean	S.D.	Rounded N	Mean	S.D.	Rounded N
Years of Education	13	2.68	91,000	13.5	2.58	40,000	13	2.63	17,000	14	2.7	7,400
High School or More	0.81	0.39	91,000	0.86	0.35	40,000	0.82	0.38	17,000	0.90	0.29	7,400
Some College or More	0.44	0.50	91,000	0.53	0.50	40,000	0.43	0.50	17,000	0.60	0.49	7,400
College or More	0.20	0.40	91,000	0.26	0.44	40,000	0.2	0.40	17,000	0.34	0.47	7,400

Appendix

Figure A1

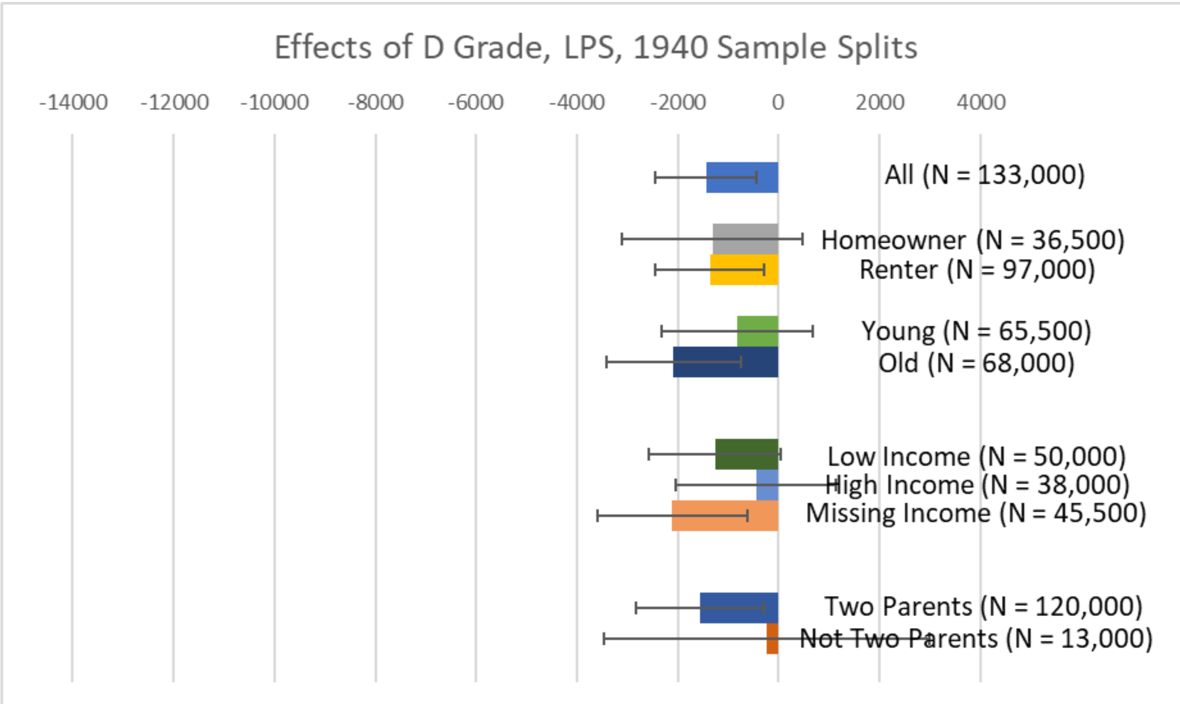


Figure A2: