Black High Schools in the Segregated U.S. South*

Soyoung Han Melvin Stephens, Jr.

March 2026

Abstract

Educational access for Black Americans has historically lagged far behind that of Whites. In particular, only one in six counties in the segregated US South had a Black high school in 1910. This paper examines the impact of the subsequent rapid spread of high schools over the next three decades on Black educational and economic outcomes using a newly compiled dataset of segregated Black high school openings. Leveraging variation in the timing and location of school openings, we find that high school exposure substantially increased the likelihood of attending high school and raised total years of schooling. We find that these increased educational levels significantly affected adult location/migration choices, occupational choices, and earnings.

*Han: University of Michigan, soyoungh@umich.edu; Stephens: University of Michigan and NBER, mstep@umich.edu. We thank seminar participants at the University of Michigan for helpful comments. We thanks Cindy Ma, Tevin Thomas, Maxwell Rambo, Donavin Stoops, Damiya Verma, and Melissa Lewis for excellent research assistance. The authors have no conflicts to disclose.

1. Introduction

Historically, access to educational opportunities for Black Americans has lagged behind those for Whites. The Freedman's Bureau along with Northern philanthropists worked to provide elementary schools for emancipated slaves. While state constitutions were rewritten after the Civil War to include funding for public schools, elementary schools spread quite slowly for Blacks. The well-known Rosenwald Fund built new schools for Blacks, particularly for rural areas, starting in late 1910s through the early part of the Great Depression, increasing rural black educational attainment and migration to the North (Aaronson and Mazumder, 2011).

While the United States experienced a great expansion of secondary schooling from 1910-1940, secondary school access for the Southern US states lagged (Goldin, 1998). Nonetheless, the proliferation of White high schools in the South was quite rapid during this same period. For Southern Blacks, the delay was far greater. Only 2.8% of high school age Blacks were enrolled in a high school in 1910 with over two-thirds of enrollment in private secondary schools (Anderson, 1988). In fact, only 64 public high schools were available for Southern Blacks in 1916 (Redcay, 1935a). However, through the combination of philanthropy and local interests, Black public secondary school access expanded rapidly and by 1933 there were more than 2000 such schools (Redcay, 1935a).

In this paper, we examine the impact of the expansion of Black high schools in the segregated Southern states on educational outcomes (completed schooling and high school attendance) and economic outcomes as adults (e.g., earnings, occupation, and mobility to the North). To do so, we compile a dataset of segregated Black high schools between 1904 and 1950, including information on openings and closings, as well as where they are located in thirteen states which mandated school racial segregation. We examine the impact of high school access using linked complete count Census data for Black males youth in the 1910-1930 Census which allow us to determine high school exposure by an individual's location

when they are growing up.

We identify the impact of being exposed to a high school by leveraging variation in the timing of high school openings. An important consideration that we document below is that school transportation is not available for the vast majority of Black (and most White) students during the period we examine. Given this limitation, the opening of a high school in a child's county may not necessarily mean that a secondary education is readily accessible. We instead measure access based on the Minor Civil Division (MCD) where the individual grew up. Information on MCDs, which are political sub-areas of counties such as election districts, is available in the restricted access version of the complete count Census data. We account for changes in MCD definitions over time by creating time-invariant MCDs using details of MCD changes found in comptemporaneous Census publications. Combining these definitions with 1940 MCD maps and our Black high schools dataset, we construct measures of high school access for those who grew up in these MCDs at different times.

In addition to better approximating school access through the use of MCD level information, this approach also allows us to account for potentially important school quality improvements both across and within states. Card and Krueger (1992) show that 15%-20% of the reduction in Black-White wage gap is due to the improvement in relative Black school quality between states between 1960 and 1980. Carruthers and Wanamaker (2017) find that variation in school quality between counties within states can explain about 60% of the Black-White wage gap in 1940. By assigning high schools at the MCD level, we can leverage within county variation in school access and control for differential trends in school quality across counties.

In our analysis, we find that exposure to high school significantly increases the likelihood of attending high school and total years of schooling using data in the 1940 Census. We then examine the longer term impact of educational attainment on a range of adult outcomes. We find that increased educational attainment affected location/migration decisions, occupational choices, and earnings (as proxied by occupational score). We plan to extend our

analysis to include the 1950 Census as additional data becomes available.

This paper contributes to multiple literatures. First, the literature on the High School Movement has largely ignored black high schools on the grounds that there were too few. We provide new insights into the degree to which Black youth were able to access high schools in this era. Second, by leveraging the variation in the the timing of schooling, we are able to provide causal evidence on the impact of education, e.g., the returns to schooling. There is limited evidence on the returns for Blacks using standard instruments such as compulsory schooling and child labor laws. Third, this paper contributes to the literature on Black educational attainment in the segregated South. Prior papers have examined the impact of school quality and the Rosenwald schools during this era. Finally, we contribute to the literature examining the impact of expanded educational access on Black mobility during the Great Migration.

2. Historical Background

2.1 The Development of High Schools in the South

In the immediate aftermath of the Civil War, there were very few established educational institutions for Black Americans in the South. While freed Blacks had established schools in the South prior to emancipation (Anderson, 1988), the establishment of the Freedman's Bureau led to thousands of Black schools being built throughout the region (Margo, 1990). In addition, Southern state constitutions were re-written during Reconstruction and mandated public funding for schools (Anderson, 1988). However, access to school for Blacks in the South remained quite limited. When Whites regained control of the Southern state governments, school funding that was distributed on a per capita basis for all children was primarily funneled to White children (Margo, 1990). Given the view that Black economic opportunities should be limited to occupations such as farming, domestic service, and menial labor, educating Black children was not viewed as necessary.

Unsurprisingly, very few black children were enrolled in high school in the South during

the late 19th Century. Among high school age children in the South, 0.4% of Blacks were enrolled in secondary schools in 1890 while the corresponding rate for Whites was 4.0% enrollment rate for Whites (Anderson, 1988). Moreover, there were very few public secondary schools available for Blacks at this time as over two-thirds of Black high school enrollment was in private schools.

However, while secondary education would begin to rapidly expand throughout the US during the first part of the twentieth century (Goldin, 1998), prospects for Blacks to enjoy a similar growth in secondary school access had already been dashed before the end of the nineteenth century. Coming on the heels of the "separate but equal" decision in *Plessy v. Ferguson* (1896), the Richmond County Board of Education in Georgia decided to close the only Black public high school in the state under the auspices of needing to fund more Black elementary schools. Since the school board maintained two high schools for Whites, one for boys and one for girls, a lawsuit was brought seeking to bring back the Black high school. However, in *Cumming v. Richmond County Board of Education* (1899) the unanimous decision of the Supreme Court, which was written by Justice Harlan who was the lone dissenter in *Plessy v. Ferguson*, concluded that such school funding decisions were within the purview of the state (Kluger, 1975).

As noted by Goldin (1998), the expansion of secondary schooling in the South lagged behind the rest of the US in terms of enrollment rates. However, the process of expanding access to high schools in the South began soon after the new century began. In 1905, the General Education Board, which was started with funding from John D. Rockefeller, began funding a professor of secondary education in each southern state. As part of their position, these professors were expected to advocate for the construction of high schools and work with local authorities to build enthusiasm for their construction (Anderson, 1988).

During this same period, many southern states also passed county high school laws. These laws required each county to maintain at least one high school. As a consequence of these actions, White high schools began to rapidly develop in the south. For example, between

1904 and 1916, Georgia moved from having 4 four-year White high schools to having 122 such schools (Anderson, 1988). While these high schools primarily were initially located in urban areas, the spread to include rural locations continued over the ensuing decades.

All these efforts, however, were aimed at providing high schools for white children but not for Blacks. For example, in the 1911-13 Biennial Report of the Superintendent of Public Instruction of Kentucky, the state High School Supervisor writes (p.363):

"County High School Law nowhere definitely states ... must organize high schools for colored pupils, but the State Common School Law definitely specifies that no distinction shall be made between the opportunities offered the white and colored children, save and except in the single fact that they must be conducted in separate institutions."

He further notes that whether there will ever be a demand to provide secondary schools for Black children "is a condition which the future alone will develop."

2.2 The Slater Fund and County Training Schools

The Slater Fund supported Black educational institutions in the South. Founded in 1882, the Fund was initially focused on supporting Black colleges and the production of Black teachers to support the growing demand for educating Black children (Redcay, 1935a). In the second decade of the twentieth century, the Fund turned its attention to the need for Black secondary schools. The lack of secondary schools meant that there was a very limited pipeline of students to attend Black colleges and normal schools which in turn meant the potential supply of new Black teachers was rather small (Redcay, 1935a).

One reason for the lack of Black secondary schools was the resistance of Whites to providing classical education to Blacks (Anderson, 1988). However, the provision of industrial skills was acceptable as it would not disrupt the economic and social order by potentially providing Blacks with skills that could upset the economic and social order of the South. Northern philanthropists were willing to follow the "Hampton-Tuskegee" model, which was to provide Blacks with an industrial education as opposed to a classical education.

The Slater Fund initiated County Training Schools (CTS) in 1911. These schools were designed to address a number of issues raised above (Redcay, 1935a). While they began as elementary schools, the plan was that they would eventually offer a few years of secondary education. The curriculum would focus on industrial and agricultural skills. In addition, the CTS would also provide teaching training to help increase the supply of available teachers, including in rural areas.

To support County Training Schools, the Slater Fund proposed a cooperative model between the school districts and the Fund. The CTS were to be public schools that were part of the local school system (Redcay, 1935a). The Slater Fund would contribute annually while local school districts also committed to annual funding of the schools. Many of the CTS were in rural areas and would eventually become four-year high schools (Redcay, 1935a). The Rosenwald Fund, which began supporting rural elementary schools soon after, would also use a cooperative funding model (Aaronson and Mazumder, 2011).

2.3 Growth of Southern Black High Schools: 1910-1930

While Southern White high schools began to rapidly increase around 1910, the spread of Black high schools lagged behind. Most Southern cities did not have Black high schools in 1916. Anderson (1988) argues that as Blacks began to migrate to the North and to the urban South from rural areas, Black high schools were built as a response to the growing numbers of idle Black teens in Southern cities. Unlike with the CTS, these high schools appear to have offered classical education. Thus, while Black secondary schools began to spread, Northern philanthropists were not that involved with these schools, at least not until the late 1920s (Anderson, 1988).

Thus, Black secondary schools spread through the South both due to County Training Schools and through more traditional high schools. The Slater Fund began by funding four CTS in 1911. Over time the number of CTS spread such that hundreds of schools annually received support from the Slater Fund. At their peak in 1931, 390 CTS received

annual support from the Slater Fund (Redcay, 1935a). As noted above, a goal of the Slater Fund was that the CTS would eventually offer multiple secondary grades, even four years of high school. Once schools became accredited four-year schools, however, they would no longer receive support from the Fund. By 1934, the Fund had supported 612 CTS since the program's inception with these schools spread across 517 counties.

During more than two decades spanning 1910-1934, Southern high school enrollment rates increased dramatically for Blacks and Whites. For Blacks, the enrollment rate of high school age children grew from 2.8% in 1910 to 18% in 1934. While there were 64 public Black high schools in the South in 1916 (Jones, 1916), there were over 2,000 such schools by 1934 (Redcay, 1935a). Interestingly, two-thirds of these high schools were in rural areas. Of these 2,000 public high schools, over 800 were four-year high schools and, of these, roughly half had received support from the Slater Fund at some point (Redcay, 1935a). While this growth in Black secondary school access was dramatic, it was dwarfed by the contemporaneous high school enrollment growth for White children, which rose from 10.1% to 54% (Anderson, 1988).

3. Data

In this section, we discuss our database of Black high school openings and closings along with locations, our approach for linking Minor Civil Divisions across Census years, and the complete count Census data used in our analysis.

3.1 High School Openings and Locations

The basis of our high school database are national lists of public and private Black secondary schools collected at five points in time between 1904 and 1949. We use the 1904-5 Report of the Commissioner of Education, Volume 2, the 1916 Bureau of Education bulletin titled Negro Education, the 1925-26 Statistics of the Education of the Negro Race, the 1931 Negro Year Book, the 1935 Public Secondary Schools for Negroes in the Southern States of the

United States, and the 1949 Directory of Secondary Schools in the United States. Most of these data collection efforts were published by the US Bureau of Education except for the 1931 Negro Year Book (published by the Tuskegee Institute) and the the 1935 Public Secondary Schools for Negroes (published by the Slater Fund).

Relatedly, the US Bureau of Education began publishing lists of accredited high schools in the 1920s. However, many high schools were not initially accredited which is particularly true for Black schools that also lagged behind in the accreditation process. Thus, only focusing on accredited schools would dramatically understate the extent to which Black children had access to secondary schooling. As such, we expand our data collection efforts beyond accredited schools.

Building on these lists, we then determined school opening and closing dates, as well as school locations (specifically, town or city) using additional data sources. First, we identify high schools through lists of high schools and/or high school principals that appear in state education directories. However, these directories tended to exclude Black schools, especially in earlier years. Second, we used state education reports which intermittently reported lists of high schools. Third, we relied on web searches which produced information on openings, closings, and locations of Black schools through a variety of means including newspaper articles, historical societies, and historical markers.

Determining the exact opening (and closing) dates of schools proved to be difficult as many have closed (or, following school desegregation, became junior high schools or elementary schools), changed names (especially when a new building was built), or had the same name as the local White high school, among other challenges. In some cases, the sparsity of schools for Blacks proved useful. E.g., since many counties had at most one Black high school, we could also use county enrollment data by grade in the state education reports to determine when a high school opened in the county.

Another issue that we address is the determination of when a school is a high school. As noted, many County Training Schools originally only offered elementary school grades before

eventually becoming high schools. Similarly, we encountered other schools that originally only offered elementary school grades before eventually offering some secondary school before providing four years of school. We designated a school as a high school based upon when it first offered four years of secondary education. From the perspective of when a child could first enter the school and graduate in four years, we set the school opening as four years prior to the year that the school first graduated students or first had positive enrollment in the fourth year of high school.

Our final list contains both public and private Black high schools that were open between 1904 and 1950, inclusive. Schools that opened earlier this period or that remained open after this period are also in our database (although we may not have an exact opening or closing date for these schools). We compiled these schools for thirteen states: Alabama, Arkansas, Delaware, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia.[1]

To determine a school's location, we primarily rely on data reported in national and state publications, which provided the town or city for each school and usually, but not always, the county as well. We assign each school to a geographic area based on the centroid of the reported city where the school is located. For cities that contain multiple geographic areas, such as Birmingham, AL and Atlanta, GA, we identify the exact street address or geographic coordinates of each school to assign the school to the correct geographic area.[2]

As a point of comparison, we also collect data on the availability of White high schools, as close in time as possible to each of the Census years between 1910 and 1940, inclusive. We are only able to produce complete lists of White high schools for four states across all four Census years: Kentucky, Maryland, North Carolina, and West Virginia. We are, however, able to create lists of accredited schools for all thirteen states which we can use to assess

[1]We have also collected high schools for Florida but we do not use data from this state in our analysis due to a relatively high level of geographic boundary changes during this period.

[2]In the next subsection we discuss the main geographic areas we use which are Minor Civil Divisions. Larger cities tended to have at least two such areas but some had many more. E.g., Birmingham, AL contained twelve such areas.

whether the trends in these four states are similar to those in all states, at least among accredited White schools.

Figure 1 shows the trends in public and private high schools in our thirteen state sample from 1904 through 1950. Consistent with this historical record, there are more private than public black high schools at the outset of the 20th Century. Black public high schools begin increasing in the 1910s and then much more rapidly in the 1920s. The number of private black high schools remain relatively constant until the 1930s when the combination of a growing number public schools along with the Great Depression leads to a decline in these schools.

Trends in the availability of Black and White high schools across counties in our sample are shown in Tables 1 and 2. Table 1 shows the share of children aged 14-17 that have at least one high school in their county. For Black children, we see that less than half have a high school in their county by 1920 and still not all have one by 1940. When looking at White students in the four states with complete high school lists, we see that the majority have a high school in their county by 1910 and virtually all have a high school available by 1920.[3]

Table 2 examines high school availability along the intensive margin. The Table shows the average number of high schools in counties among those that have at least one school available. We observe that the number of schools per county available is roughly constant for Black students which suggests that much of the growth of Black schools is along the extensive margin of having a school in your county. For White students, we observe a near tripling in the number of schools per county through 1930 and a slight down turn in this number in 1940. This subsequent decline is consistent with the onset of school consolidation during this period where school transportation began to gain traction and smaller local high schools were replaced with larger, centralized schools.

[3]If we focus only on White accredited schools (not shown here), the shares of White students with an accredited school nearby is very similar to our complete thirteen state sample as well as in the four state subsample with complete White high school listings.

Using our dataset of Black high schools, we can provide some initial insight into how school access affected Black high school enrollment. For eight of our thirteen states (excluding Arkansas and Virginia), we are able to obtain Black county-level enrollment data for high school grades from state education reports for three Census years, 1920-1940.[4] We construct Black county high school enrollment rates by dividing county high school enrollment by the number of Black children ages 14-17.

Figure 2 plots enrollment rates by year based on when the county first opened a high school: before 1920, between 1920 and 1930, between 1930 and 1940, and after 1940.[5] As shown in the Figure, outside of the counties that already have high schools in 1920, there is very little enrollment in the remaining counties. As we noted above, some schools may offer some high school grades before they are counted as a high school based on our definition of offering four years of high school. As anticipated, there are large enrollment gains between 1920 and 1930 for the counties that first open a high school between these two years while the counties without high schools show little enrollment growth. Similarly, the counties that have a first Black high school open between 1930 and 1940 show a large enrollment gain while there is only a slight increase in enrollment among those counties that remain without a high school.

3.2 Minor Civil Divisions: Linking County Sub-areas Across Time

Studies that examine the impact of schooling requirements and school quality typically use measures defined at the state or county level. However, because school transportation was quite limited during this period, especially for Black children, measuring schooling access at these higher geographic levels may not adequately capture the available educational opportunities. E.g., many of the initial compulsory schooling laws in Southern states did not impose attendance requirements on children who lived more than two or three miles from

[4]These states are Alabama, Delaware, Georgia, Kentucky, Maryland, North Carolina, Tennessee, and West Virginia.

[5]Figure 2 shows public school enrollment although the measure of school availability includes both public and private schools.

the nearest school.[6]

Figure 3 presents race-specific transportation rates for all children in our sample. Most states do not report the number of transported students prior to the 1930.[7] As the Figure shows, very few Black students were transported around 1930. To some extent, these numbers understate the racial differences in the availability of transportation. E.g., while South Carolina had roughly 200,000 each of white and black elementary school children in 1940, the state funded 1,644 buses for white students but only 8 buses for black students. Data on transportation rates for high schools is available for some states as shown in Figure 4. Although transportation rates were somewhat higher for secondary school students, these rates are still relatively low across most states throughout the entire period. More importantly for our analysis, very few Black high school students were transported prior to the 1930s.

To better match children to the available high schools, we measure school access at the level of the Minor Civil Division (MCD). MCDs are county sub-areas although their precise definition varies by state. E.g., MCDs are election districts in Alabama while in Arkansas MCDs are townships. The enumeration districts set by the Census Bureau do not cross MCD boundaries. In the least populous areas, MCDs contain a single enumeration district. Importantly, MCDs are considerably smaller than counties. Figures 5 and 6 show the 1940 county and MCD maps, respectively, for Alabama. While the average county in our sample is 452 square miles in 1940, the average MCD in the same states is 44 square miles.

One issue that arises with the use of MCDs is that their definitions may change over time. To address boundary changes, we use information on adjustments to MCD boundaries from Census publications to construct time-consistent MCDs between 1910 and 1940. For example, Figure 7 shows the 1940 MCDs for Marion County, AL. In 1922, precinct 19 was

[6]Attendance exemptions varied across states: students residing more than 2.5 miles from the nearest school were exempt in Alabama, more than 2 miles in Delaware and Florida, and more than 3 miles in Georgia. (Gateway to Global Aging Data, 2025)

[7]State education reports do contain information on transportation dollars spent in earlier years although it is difficult to infer transportation access from this information.

created from parts of precincts 6 and 7. Thus, we create a single time-consistent MCD that combines precincts 6 and 7 in the 1910 and 1920 Censuses and combines precincts 6, 7, and 19 in 1930 and 1940 Censuses as shown in Figure 8. While the original 1940 MCDs average 49.56 square miles, our time-consistent MCDs average 54.73 square miles. Finally, some counties were entirely redistricted one or more times during this period. We exclude these counties from our analysis.

3.3 Complete Count Census

We use individual-level data from the 1910-1940 complete count US Census provided by IPUMS (Ruggles et al., 2025). The data include geographic details such as current state and county of residence as well as demographic information such as race and age. Through a restricted-use agreement, we also have information on the household's MCD of residence. The 1940 Census also includes detailed information on each individual's highest grade completed.

Our analysis focuses on Black males because linkage rates between Censuses are substantially lower for Black females. We construct a pooled sample of individuals aged 8 to 17 who were enumerated in the 1910, 1920, or 1930 U.S. Censuses in thirteen Southern states.[8] These individuals are then linked forward to the 1940 Census to observe later outcomes. [9] The linking of Censuses is performed using the Census Tree Project, and we include all individuals who were successfully matched using one or more linkage methods (Buckles et al., 2025; Price et al., 2021). To ensure consistency, we restrict the sample to individuals who are also recorded as Black males in 1940. Linkage rates decline with the temporal distance between censuses: 33 percent for 1910–1940, 38 percent for 1920–1940, and 50 percent for 1930–1940.[10] To address potential selection arising from differential linkage probabilities, we construct inverse propensity-score linkage weights following Bailey et al. (2020). Finally,

[8]These states are Alabama, Arkansas, Delaware, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia.

[9]For our analysis, our sample is limited to ages 18-47 in 1940. In future iterations, we plan to incorporate the 1950 Census and adjust our birth cohorts (e.g., not use 18 year olds) when examining adult labor market outcomes.

[10]The analogous linkage rates for Black females are 8.8, 11.5, and 28.1 percent.

counties with 1910, 1920, or 1930 MCDs that could not be matched to 1940 MCDs are excluded from the analysis.

We define exposure to high school based on whether a high school was available in an individual's MCD or county of residence when they were between the ages of 14 to 17, inclusive. The MCD or county that we use to determine high school exposure is determined by the individual's location when they appear in the earlier Census when they were between ages 8 and 17. For example, if a high school first opened in 1910, individuals who were 17 years old in 1910 are classified as having zero years of exposure, those who were 16 years old are assigned one year of exposure, those who were 15 years old would have had two years of exposure, etc.[11] While our example is based on a school opening in a Census year, we use all school openings in our analysis. In our results shown below, we only use a binary measure of schooling exposure defined as whether an individual has any exposure to a high school.

4. Research Design

This section describes the research design employed to estimate the effects of high school openings on the likelihood of attending high school and total years of schooling.

4.1 Impact on Educational Outcomes

We leverage variation in the timing of high school openings across locations to examine the impact of high school access on educational outcomes. For our baseline models, we estimate expanded difference-in-differences specifications based on the equation

$$y_{igy} = \beta_0 + \beta_1 \, 1(High\ School\ Access)_{igy} + \mathbf{X_{igy}}\theta + \alpha_g + \gamma_{g'y} + \epsilon_{igy} \tag{1}$$

[11] We assign school exposure in this way for two reasons. First, the Census collects age not year of birth. Since the Census is enumerated as of April 1, 3/4ths of individuals will not have passed their birthday in that calendar years. Second, school openings usually refer to schools opening in the fall. In our example, someone who is 17 in the 1910 Census and experiences a school opening in the same year is most likely turning 18 when the school opens in the fall of that year.

where y_{igy} is the educational outcome for individual i, born in year y, who grew up in geographic unit g, observed in 1940. The binary indicator $High\ School\ Access_{igy}$ equals one if the individual has at least one year of exposure to high school, as defined in Section 3.3. We include MCD fixed effects, α_g, which capture time-invariant characteristics of the MCD where the individual grew up. We also include year of birth fixed effects that vary by higher-level geography, $\gamma_{g'y}$, to account for differential changes across birth cohorts in local factors such as school quality. When high school access is measured at the MCD level, the $\gamma_{g'y}$ corresponds to county-by-birth-year fixed effects; when access is measured at the county level, the $\gamma_{g'y}$ are state-by-birth-year fixed effects. Finally, in some specifications we add demographic information from when the individual is observed between the ages of 8 and 17: if they lived on a farm, if they lived in an urban area, whether they resided with their mother, and whether they resided with their father.

In our analysis, we examine the impact of high school access on multiple educational outcomes. The first outcome is a binary indicator for whether the individual completed at least one year of high school. The second outcome is the total years of schooling that the individual completed. Finally, we extend the use of binary indicator outcomes to cover all years of schooling ranging from one to sixteen years in order to look at how high school access impacts the CDF of completed education.

We estimate two sets of difference-in-differences (DD) models. The first employs a two-way fixed effects (TWFE) specification. Following the suggestion by Sun and Abraham (2021), we exclude geographic areas whose first high school opened before 1910. Because our setting involves staggered adoption of high schools, the presence of treatment effect heterogeneity across treatment cohorts can induce bias in the TWFE estimates (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2020; Sun and Abraham, 2021; Borusyak et al., 2024). To address this concern, we also report DD estimates using the estimator proposed by de Chaisemartin and D'Haultfœuille (2024) (dCDH), which is robust to heterogeneous treatment effects with staggered treatment timing.

Standard errors are clustered at the MCD level throughout.

4.2 Impact of Education on Adult Outcomes

In addition, we examine the impact of additional schooling gained through access to a high school impacts adult outcomes. To do so, we estimate equations of the form

$$y_{igy} = \delta_0 + \delta_1 \, Educ_{igy} + \mathbf{X_{igy}}\omega + \lambda_g + \eta_{g'y} + u_{igy} \tag{2}$$

where y_{igy} is an adult outcome for individual i, born in year y, who grew up in geographic unit g, observed in 1940. We include MCD fixed effects, λ_g, which capture time-invariant characteristics of the MCD where the individual grew up. We also include year of birth fixed effects that vary by higher-level geography, $\eta_{g'y}$, to account for differential changes across birth cohorts in local factors such as school quality. We again include additional demographic controls, $\mathbf{X_{igy}}$. We are primarily interested in the coefficient on years of schooling $Educ_{igy}$. For these specifications, we use the binary indicator $High\ School\ Access_{igy}$ to instrument for years of schooling. In addition, for our adult outcomes other than educational attainment, we restrict the analysis to cohorts observed in the 1910 and 1920 Censuses.

The first set of adult outcomes we examine involve location choices. We examine whether increased educational attainment affected the propensity to live on a farm or to reside in an urban area. In addition, we examine whether increased educational attainment due to high school access may have contributed to the Great Migration by examining the effect on the probability of moving North.

The second set of results involve labor market outcomes. In particular, we examine whether increased educational attainment impacted occupational choice. We also measure the impact on the individual's occupational score which is a proxy for earnings.

5. Results

5.1 Effects on Educational Attainment

Table 3 reports estimates of the impact of high school access, measured at the MCD level, on educational outcomes. The first four columns report results using the completion of any high school. The results show estimates of (1) using both TWFE and the dCDH estimators. For each estimation method, we show results with and without individual demographic covariates. While the reported estimates for the coefficient on $High\ School\ Access_{igy}$ are statistically significant across all the specificiations, the results using the dCDH method are slightly smaller and including covariates also yields slightly smaller estimates. Focusing on the dCDH estimates which include covariates, we find that Black males exposed to a high school were 3.5 percentage points more likely to have completed at least one year of high school, a 29 percent increase relative to the mean of 12.2%.

In the last four columns of Table 3, we report estimates using total years of schooling as the outcome in (1). Again, we find that the estimates of the coefficient on $High\ School\ Access_{igy}$ are significant in all specifications. Our dCDH estimates including covariates yields that having local access to a high school increases total years of schooling increase by 0.24 years of schooling, representing a roughly 5 percent gain.

Table 4 presents analogous estimates but where we have measured exposure to high school at the county level. As we noted above, since state supported school transportation was essentially unavailable for Black students prior to the 1930s, the high school access measured at the county level likely overstates the extent to which Black children could readily access a secondary education. Consistent with this observation, we find much smaller effects when access is measured at the county level as opposed to the MCD level, although we continue to find that high school access led to increased schooling.

We show the results from multiple robustness checks, using alternative specifications of equation (1), in Appendix A. When high school access is measured at the MCD level, we

find that using state by birth year fixed effects instead of county by year fixed effects for $\gamma_{g'y}$ has very little impact on our estimates (Table A.1). However, if we use nationally defined birth year fixed effects, our estimates increase by approximately 10% - 30%, suggesting that local differential changes, such as school quality, may influence educational results during this period. When high school access is defined at the county-level, we also find larger point estimates when using nationally defined birth year fixed effects (Table 4).

Finally, we estimate the impact of high school access on one minus the CDF of education. For each year Y of schooling between one and sixteen years, inclusive, we estimate equation (1) where the outcome is whether years of schooling exceed Y. We would expect high school access to primarily affect years of high school attendance. However, the availability of a high school might encourage students who would have otherwise dropped out of elementary school to continue to high school. In addition, as noted earlier, many high schools were union schools meaning they offered both elementary and high school grades which also may have impacted the completion of earlier grades.

Figure 9 presents these estimates in two complementary ways. The first approach, shown by the solid line in the figure, is coefficient on $High\ School\ Access_{igy}$ which shows the percentage point increase in education at each level of completed schooling. The results show no impact of high school access on early elementary school grades and, relatedly, little impact beyond on high school grades. While we find significant effects at all high school years, the largest percentage point impacts are in the two years immediately prior to high school.

However, as shown in the first column of Appendix Table A.3, the share of Black men who have completed a given level of schooling or more falls sharply as schooling levels increase. An implication is that the same percentage point increase might have quite different impacts on percentage increase in schooling across the education distribution.

The dashed line in Figure 9 measures the impact of high school access relative to the share of individuals who have completed a given level of education. Viewed in this way, the

largest relative increases in educational attainment, due to high school access, occur during the high school grades.

5.2 Impact of Education on Adult Outcomes

We estimate the impact of education on adult location choices using equation (2) where we use local high school access at the MCD level as an instrumental variable for educational attainment. The results are shown in Table 5. We find that additional schooling induced by having a high school in one's MCD reduces the probability of living on a farm and increases the chance of residing in an urban area. For both of these outcomes, the OLS and IV estimates are the same sign although the magnitude of the IV estimate is noticeably larger in magnitude. Interestingly, the estimates when using Northern migration as the outcome yield opposite signed, and statistically significant, results between the OLS and IV estimators.

Table 6 estimates the impact on education on occupational choice. For this analysis, we limit our sample to working men who are defined as individuals with positive income and positive working hours in the previous year. Our IV estimates show that increased educational attainment lowered the likelihood of being a farmer while raising the probability of working as an operative and as a teacher. Finally, we examine the impact of educational attainment on occupational score which is a measure of earnings potential based on one's occupation. Using this outcome, we find a positive and significant estimate indicating that the movement of Black workers across occupations, which was induced by having high school access, led to higher earnings.

6. Conclusion

Using a newly compiled dataset on the openings of high schools for Black children in the South, we examine the impact of high school access on long-run outcomes such as completed schooling, earnigns, and migration. Our analysis of thirteen states using linked data from

the 1910-1940 complete count Censuses shows that exposure to high schools increased Black educational attainment. We find that these changes in educational attainment impacted location choice with fewer men living on farmers and more living in urban settings. Perhaps surprisingly, we find that these education increases reduced Northern migratino. Finally, we find that these education increases impacted the occupational distribution of Black men which resulted in them moving, on average, to higher paying jobs. We plan to extend this analysis to the 1950 Census once better data is released and to extend our analysis to examine outcomes for women. Our study contributes to multiple literatures regarding the impact of education on Black economic progress and the widespread adoption of secondary education in the US.

Tables and Figures

Table 1: % of 14-17 Year Olds with a High School in Their County

	1910	1920	1930	1940
Blacks	32.1	42.4	75.5	90.5
Whites*	80.1	98.7	98.6	100

Table 2: Number of High Schools in Counties with High Schools

	1910	1920	1930	1940
Blacks	1.4	1.4	1.5	1.7
Whites*	2.3	4.4	6.7	6.3

*Note: Complete White high school list only available in four states

Table 3: Impact of High School Access at the MCD Level on Education

	Any High School				Total Years of Schooling			
	TWFE		dCDH		TWFE		dCDH	
High School Access	0.055	0.048	0.038	0.035	0.353	0.292	0.235	0.200
	(0.003)	(0.003)	(0.010)	(0.010)	(0.025)	(0.025)	(0.093)	(0.094)
MCD FE	✓	✓	✓	✓	✓	✓	✓	✓
County-by-Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Covariates		✓		✓		✓		✓
Mean Y	0.122				5.22			

Table 4: Impact of High School Access at the County Level on Education

	Any High School				Total Years of Schooling			
	TWFE		dCDH		TWFE		dCDH	
High School Access	0.009	0.008	0.015	0.014	0.097	0.083	0.185	0.176
	(0.002)	(0.002)	(0.007)	(0.007)	(0.024)	(0.022)	(0.065)	(0.064)
County FE	✓	✓	✓	✓	✓	✓	✓	✓
State-by-Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Covariates		✓		✓		✓		✓
Mean Y	0.122				5.19			

Table 5: Impact of Schooling on Adult Location

	Live on a Farm		Urban Residence		Moved North	
	OLS	IV	OLS	IV	OLS	IV
Years of Schooling	-0.036	-0.096	0.043	0.084	0.034	-0.041
	(0.0004)	(0.026)	(0.0004)	(0.026)	(0.0003)	(0.023)
MCD FE	✓	✓	✓	✓	✓	✓
County-by-Cohort FE	✓	✓	✓	✓	✓	✓
Baseline Covariates	✓	✓	✓	✓	✓	✓
First Stage F		34.6		34.6		34.6
Mean Y	0.38		0.44		0.20	

Table 6: Impact of Schooling on Labor Market Outcomes for Working Men

	White Collar		Farmer		Crafts		Operative	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Years of Schooling	0.017	-0.0008	-0.031	-0.057	0.005	-0.011	0.006	0.050
	(0.0003)	(0.010)	(0.0004)	(0.021)	(0.0002)	(0.012)	(0.0003)	(0.020)
MCD FE	✓	✓	✓	✓	✓	✓	✓	✓
County-by-Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Covariates	✓	✓	✓	✓	✓	✓	✓	✓
First Stage F		33.43		33.43		33.43		33.43
Mean Y	0.04		0.23		0.05		0.17	

	Unskilled		Teacher		log(occscore)	
	OLS	IV	OLS	IV	OLS	IV
Years of Schooling	0.003	0.019	0.006	0.022	0.026	0.048
	(0.0005)	(0.024)	(0.0002)	(0.005)	(0.0003)	(0.021)
MCD FE	✓	✓	✓	✓	✓	✓
County-by-Cohort FE	✓	✓	✓	✓	✓	✓
Baseline Covariates	✓	✓	✓	✓	✓	✓
First Stage F		33.43		33.43		33.43
Mean Y	0.50		0.006		2.84	

Figure 1: Growth in Black High Schools Over Time

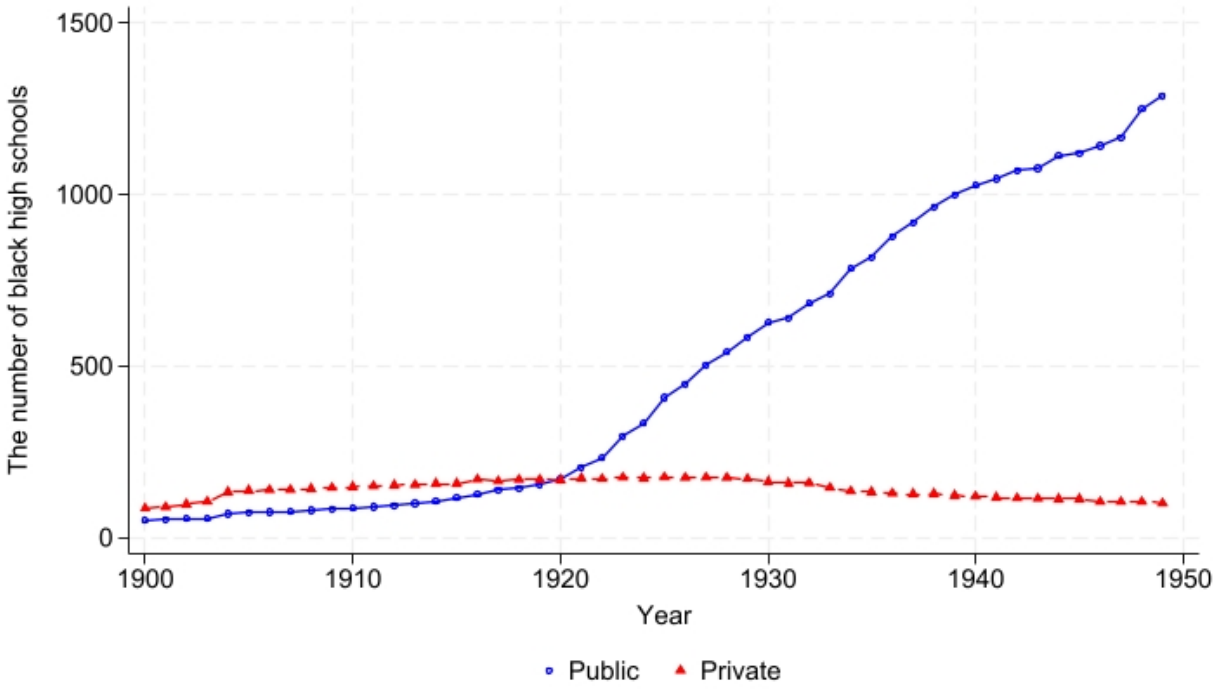


Figure 2: Impact of High School Openings on Black High School Enrollment Rates

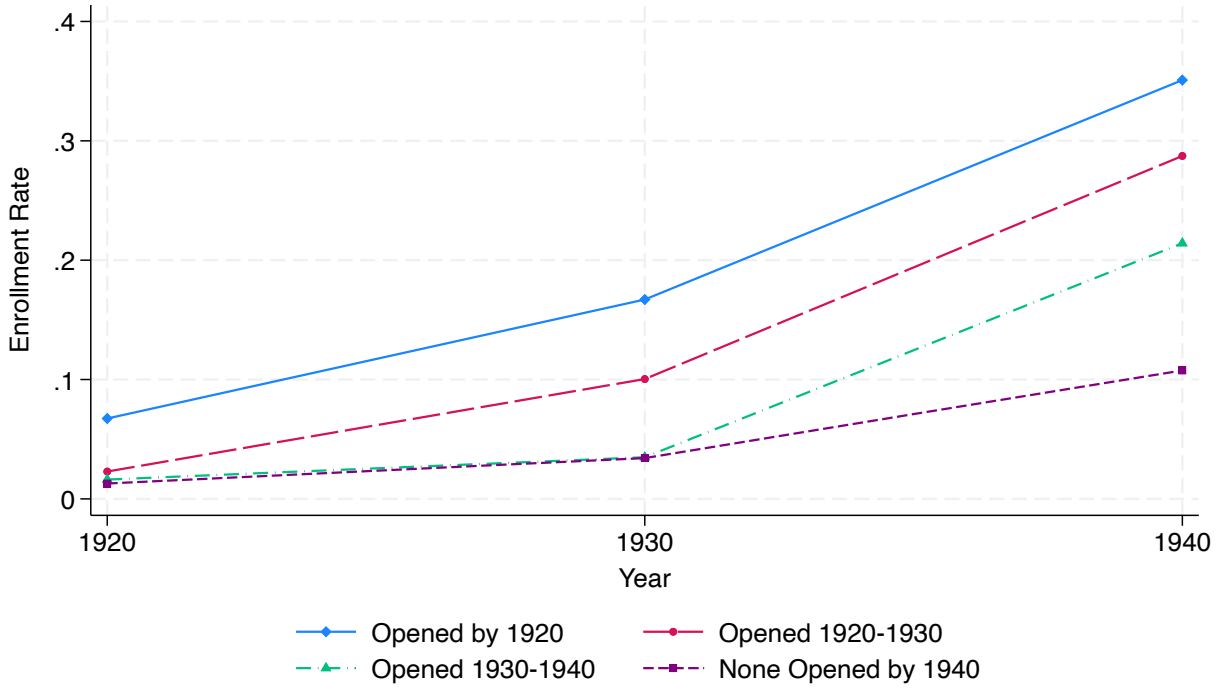


Figure 3: Share of Students Transported - All Schools

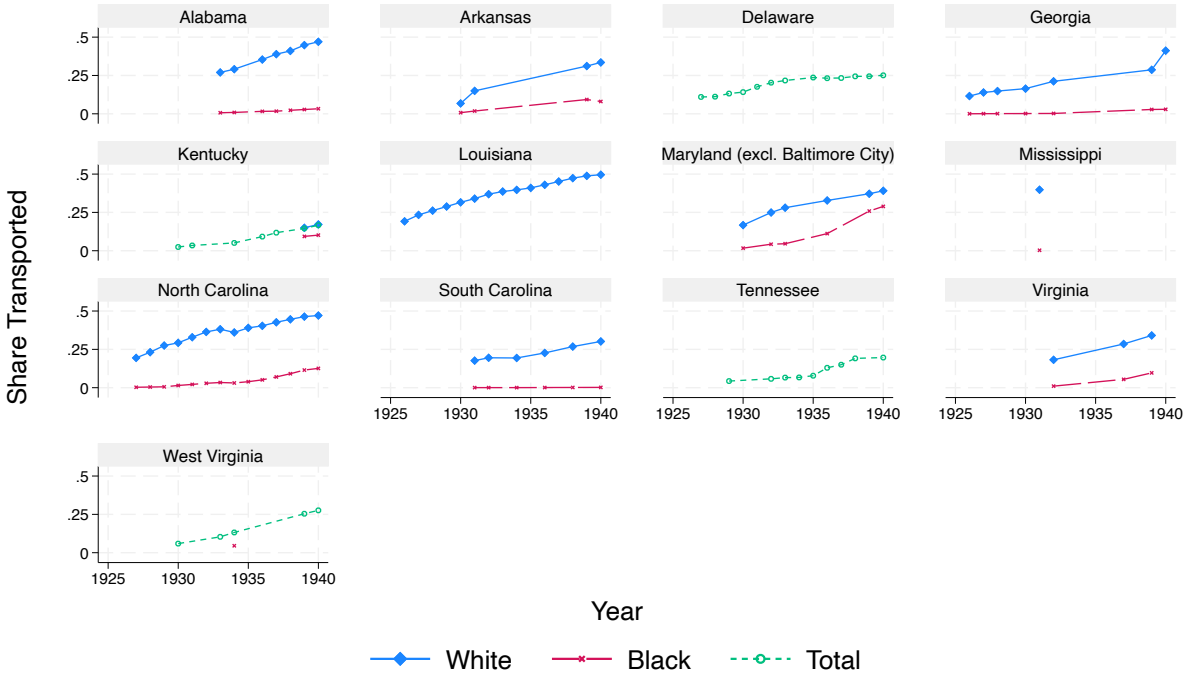


Figure 4: Share of Students Transported - High Schools

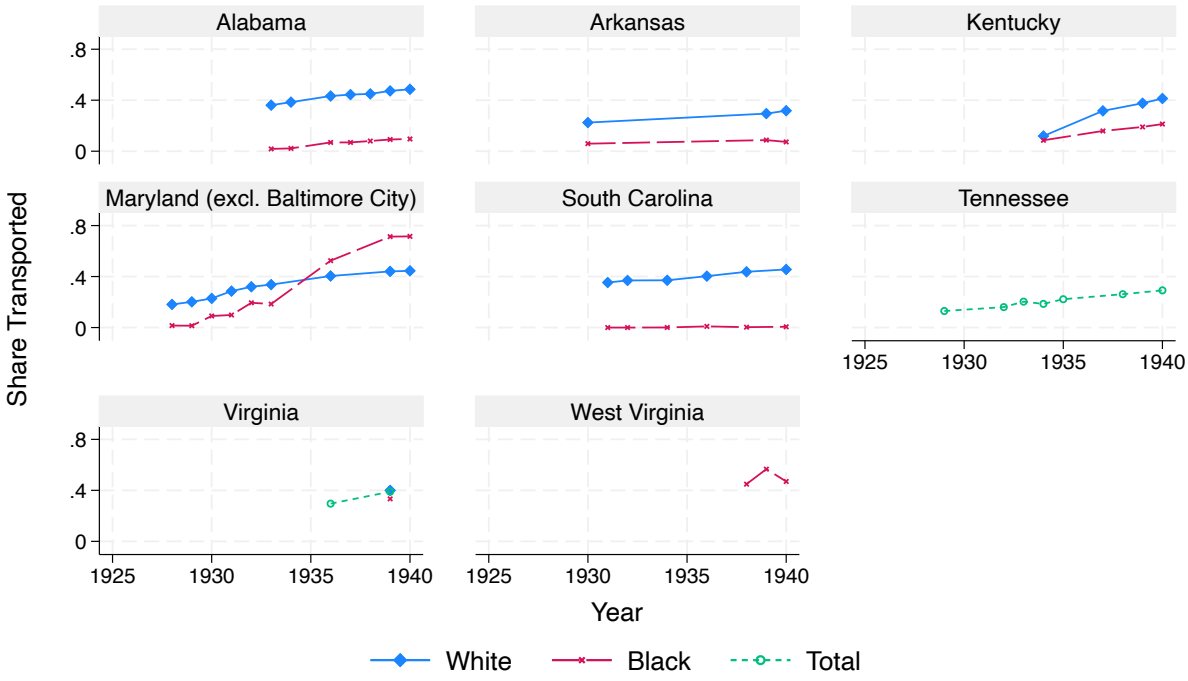


Figure 5: Alabama 1940 County Map



Figure 6: Alabama 1940 Minor Civil Division Map

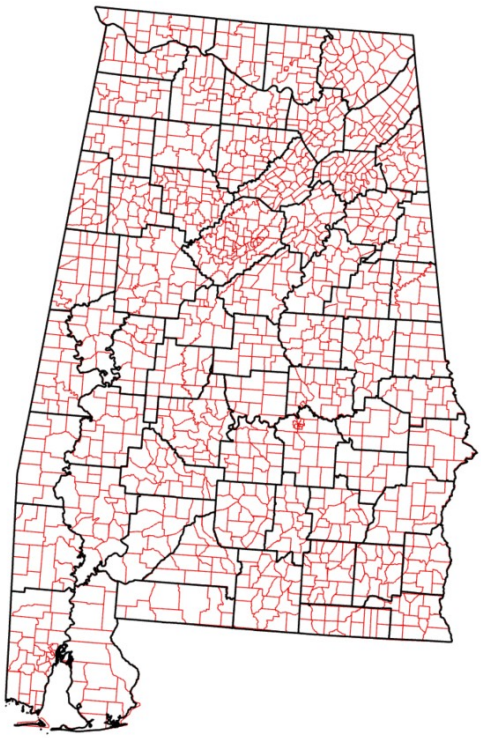


Figure 7: 1940 Marion County, AL MCD map

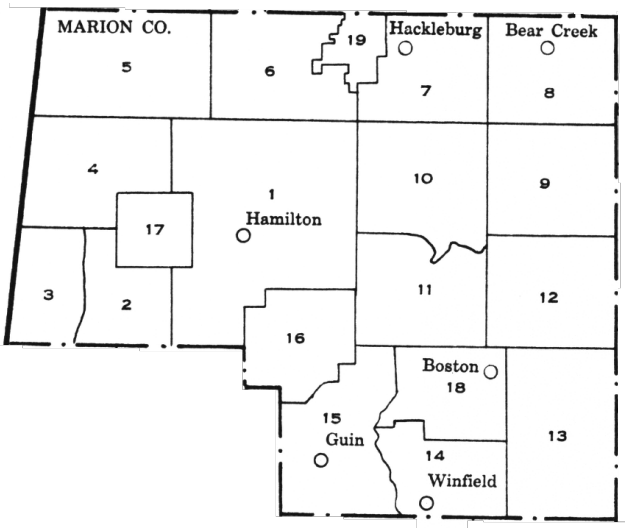


Figure 8: 1910-1940 Marion County, AL Linked MCDs map

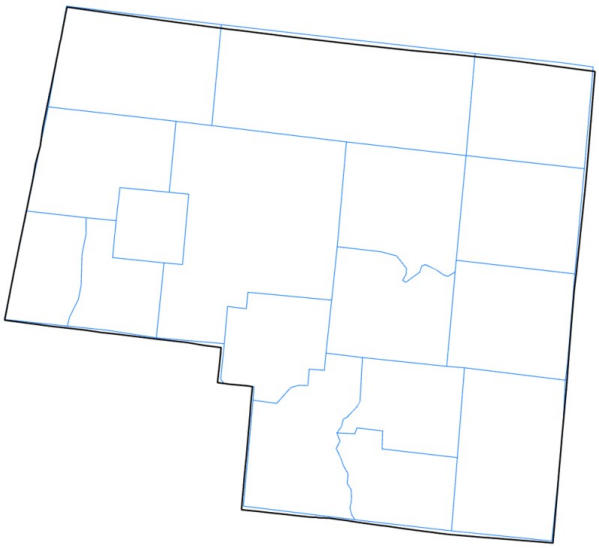
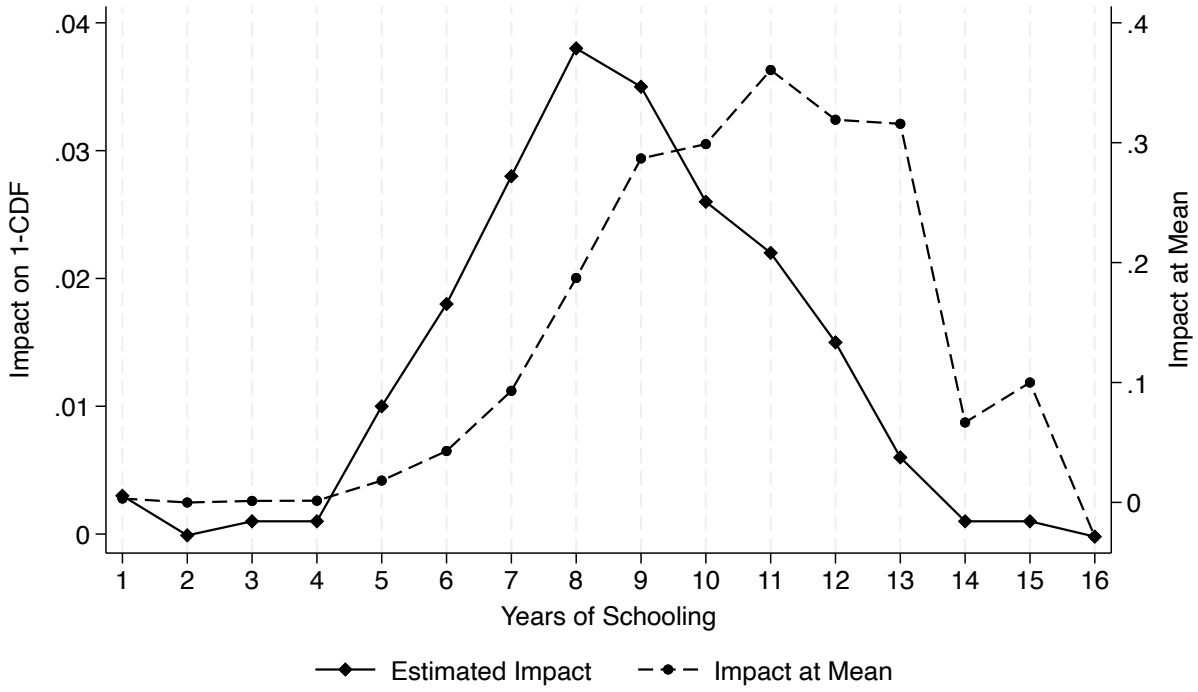


Figure 9: Impact of High School Access on 1-CDF of Education



References

Aaronson, Daniel and Bhashkar Mazumder (2011) "The Impact of Rosenwald Schools on Black Achievement," *Journal of Political Economy*, 119 (5), 821–888, http://www.jstor.org/stable/10.1086/662962.

Anderson, James D. (1988) *The Education of Blacks in the South, 1860-1935*, Chapel Hill: The University of North Carolina Press, https://research.ebsco.com/linkprocessor/plink?id=646e1b78-5bcc-3c9c-a3f9-5d71ab653122.

Bailey, Martha J., Connor Cole, Morgan Henderson, and Catherine Massey (2020) "How Well Do Automated Linking Methods Perform? Lessons from US Historical Data," *Journal of Economic Literature*, 58 (4), 997–1044, 10.1257/jel.20191526.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024) "Revisiting Event-Study Designs: Robust and Efficient Estimation," *The Review of Economic Studies*, 91 (6), 3253–3285, 10.1093/restud/rdae007.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley E.B. Wilbert (2025) "Breakthroughs in historical record linking using genealogy data: The Census Tree project," *Explorations in Economic History*, 98, 101717, https://doi.org/10.1016/j.eeh.2025.101717.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021) "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 225 (2), 200–230, https://doi.org/10.1016/j.jeconom.2020.12.001, Themed Issue: Treatment Effect 1.

Card, David and Alan B. Krueger (1992) "School Quality and Black-White Relative Earnings: A Direct Assessment*," *The Quarterly Journal of Economics*, 107 (1), 151–200, 10.2307/2118326.

Carruthers, Celeste K. and Marianne H. Wanamaker (2017) "Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap," *Journal of Labor Economics*, 35 (3), 655–696, 10.1086/690944.

de Chaisemartin, Clément and Xavier D'Haultfœuille (2020) "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110 (9), 2964–96, 10.1257/aer.20181169.

——— (2024) "Difference-in-Differences Estimators of Intertemporal Treatment Effects," *The Review of Economics and Statistics*, 10.1162/rest_a_01414.

Gateway to Global Aging Data (2025) "Gateway Policy Explorer: USA Compulsory Schooling Policy Details, 1900-2025, Version 1.2 (August 2025)," University of Southern California, Los Angeles, 10.25553/gpe.edu.ce.usa.

Goldin, Claudia (1998) "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century," *The Journal of Economic History*, 58 (2), 345–374, 10.1017/S0022050700020544.

Goodman-Bacon, Andrew (2021) "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 225 (2), 254–277, https://doi.org/10.1016/j.jeconom.2021.03.014, Themed Issue: Treatment Effect 1.

Jones, Thomas Jesse (1916) *Negro Education: A Study of the Private and Higher Schools for Colored People in the United States*, Bureau of Education Bulletin, No. 38, Washington, D.C.: U.S. Bureau of Education, Department of the Interior, `https://eric.ed.gov/?id=ED542634`.

Kluger, Richard (1975) *Simple Justice: The History of Brown v. Board of Education and Black America's Struggle for Equality*, New York: Alfred A. Knopf.

Margo, Robert A. (1990) *Race and schooling in the South, 1880-1950 : an economic history*, NBER series on long-term factors in economic development, Chicago: University of Chicago Press, 10.7208/9780226505015.

Price, Joseph, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley (2021) "Combining family history and machine learning to link historical records: The Census Tree data set," *Explorations in Economic History*, 80, 101391, https://doi.org/10.1016/j.eeh.2021.101391.

Redcay, Edward E. (1935a) *County Training Schools and Public Secondary Education for Negroes in the South*, Washington, D.C.: John F. Slater Fund, `https://archive.org/details/countytrainingsc00redcrich`, "County training schools and public secondary education for Negroes in the South" (Archive.org identifier: countytrainingsc00redcrich).

Ruggles, Steven, Sarah Flood, Matthew Sobek et al. (2025) "IPUMS USA: Version 16.0 [dataset]," 10.18128/D010.V16.0.

Sun, Liyang and Sarah Abraham (2021) "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 225 (2), 175–199, https://doi.org/10.1016/j.jeconom.2020.09.006, Themed Issue: Treatment Effect 1.

Online Appendix

A Robustness Checks on the Effects on Educational Outcomes **32**

A. Robustness Checks on the Effects on Educational Outcomes

Table A.1: Impact of High School Access at the MCD Level on Education

	Any High School							
	TWFE		dCDH		TWFE		dCDH	
High School Access	0.056	0.050	0.037	0.035	0.061	0.055	0.047	0.044
	(0.003)	(0.003)	(0.008)	(0.008)	(0.004)	(0.004)	(0.007)	(0.007)
MCD FE	✓	✓	✓	✓	✓	✓	✓	✓
State-by-Cohort FE	✓	✓	✓	✓				
Cohort FE					✓	✓	✓	✓
Baseline Covariates		✓		✓		✓		✓
Mean Y	0.123							
	Total Years of Schooling							
	TWFE		dCDH		TWFE		dCDH	
High School Access	0.364	0.305	0.251	0.222	0.41	0.351	0.329	0.299
	(0.029)	(0.029)	(0.078)	(0.078)	(0.032)	(0.032)	(0.065)	(0.065)
MCD FE	✓	✓	✓	✓	✓	✓	✓	✓
State-by-Cohort FE	✓	✓	✓	✓				
Cohort FE					✓	✓	✓	✓
Baseline Covariates		✓		✓		✓		✓
Mean Y	5.21							

Table A.2: Impact of High School Access at the County Level on Education

	Any High School				Total Years of Schooling			
	TWFE		dCdH		TWFE		dCdH	
High School Access	0.011	0.010	0.027	0.025	0.127	0.112	0.257	0.243
	(0.003)	(0.003)	(0.005)	(0.005)	(0.026)	(0.026)	(0.048)	(0.047)
County FE	✓	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Covairates		✓		✓		✓		✓
Mean Y	0.122				5.19			

Table A.3: Impact of High School Access at the MCD Level on 1−CDF Education

	Share Education>= Y	Impact of HS Access (dCDH)
Years of School:		
1 year	0.932	0.003
		(0.007)
2 years	0.901	-0.0001
		(0.008)
3 years	0.826	0.001
		(0.009)
4 years	0.711	0.001
		(0.011)
5 years	0.550	0.010
		(0.012)
6 years	0.420	0.018
		(0.013)
7 years	0.301	0.028
		(0.013)
8 years	0.203	0.038
		(0.011)
9 years	0.122	0.035
		(0.010)
10 years	0.087	0.026
		(0.009)
11 years	0.061	0.022
		(0.007)
12 years	0.047	0.015
		(0.006)
13 years	0.019	0.006
		(0.004)
14 years	0.015	0.001
		(0.003)
15 years	0.010	0.001
		(0.003)
16 years	0.007	-0.0002
		(0.002)
MCD FE		✓
County-by-Cohort FE		✓
Baseline Covariates		✓