The Geography of Black Economic Progress After Slavery*

Lukas Althoff† Hugo Reichardt‡

[Most recent version here]

August 29, 2022

Abstract

This paper studies the long-run impact of racially oppressive institutions—from slavery to Jim Crow—on Black American families until today. To do so, we link decades of U.S. census and administrative records and develop new methods to measure how long a family's ancestors were enslaved and how intensively they experienced Jim Crow after slavery. We find that a family's long-run economic progress critically depends on the location in which they were freed. Black families who were freed in states with more intensive Jim Crow institutions made less economic progress after slavery. In a border discontinuity design, we show that families freed just inside of states with a strong Jim Crow regime fared far worse than those freed just outside of those states. These gaps did not exist before Jim Crow and did not arise for white Americans. We also provide evidence that limiting Black children's access to education was a key channel.

*We thank Leah Boustan, Ellora Derenoncourt, Jeremiah Dittmar, Jonathon Hazell, Ethan Ilzetzki, Ilyana Kuziemko, Camille Landais, David Lee, Ben Moll, Suresh Naidu, Steve Redding, Ricardo Reis, Maarten de Ridder, Tianyi Wang, and Zach Ward for insightful comments. We thank all seminar and conference participants for sharing their suggestions. Tre' McMillan, Cynthia Nwankwo, and Bracklinn Williams provided excellent research assistance. This work was supported by Princeton's Program for Research on Inequality.

†Department of Economics, Princeton University. lalthoff@princeton.edu

‡Department of Economics, London School of Economics. h.a.reichardt@lse.ac.uk

1. INTRODUCTION

Throughout U.S. history, the socioeconomic status of Black Americans has remained below that of other racial groups. For example, the average Black American owns 80 percent less wealth than the average white American today (Derenoncourt et al., 2022). The disparities that Black Americans face may be in part the result of a series of racially oppressive institutions that they experienced historically.

This paper studies the long-run impact of racially oppressive institutions—from slavery to institutionalized segregation called Jim Crow—on Black American families between 1850 and 2000. Black families differ in how their ancestors experienced both slavery and Jim Crow. Their experiences were to a large degree determined by *when* they were freed—determining how long they were oppressed by slavery—and *where* they were freed—determining the intensity of Jim Crow institutions they faced after emancipation. Those two factors are closely related: due to the rapid South-West expansion of the the U.S. plantation economy in the final decades of slavery, the later a family was freed, the more likely it was to be freed in the Deep South—Alabama, Georgia, Louisiana, Mississippi, and South Carolina. 89 percent of Black families were enslaved until the Civil War (1861–1865) and most of them concentrated in the Deep South, while 11 percent of Black families were freed an average of 50 years earlier, mostly in the Upper South and North.[1] Due to low geographic mobility, the place in which a family was freed typically determined the environment in which their descendants would live for generations, including the Jim Crow institutions they faced between 1877 and 1964, such as limits to their ability to vote or the racial segregation of schools.

We overcome the empirical challenge of measuring when and where a Black family was freed from slavery—i.e., how long they were enslaved and how intensively they were exposed to Jim Crow after slavery. Our first method identifies Black families freed before the Civil War (the "Free") as those enumerated in the 1850 and 1860 censuses and identifies their descendants using historical record linkage (Abramitzky et al., 2019); everyone else is classified as descending from families freed in 1865 (the "Enslaved") because the 1850 and 1860 censuses did not record enslaved people.[2,3] Record linkage

[1] The Lower South comprises the Deep South and Arkansas, Florida, and Texas. The Upper South comprises Delaware, Washington, D.C., Kentucky, Maryland, Missouri, North Carolina, Tennessee, Virginia, and West Virginia. The North comprises all other states.

[2] We refer to Black Americans who were free before 1865 as "the Free" even though they or their ancestors may have been enslaved at some point earlier than 1865. We refer to those enslaved until 1865 as "the (formerly) Enslaved." We choose this terminology to avoid confusion engendered by the sometimes-used terms "Freemen" (Free) and "Freedmen" (formerly Enslaved). We avoid the term "slave" and capitalize "Free" and "Enslaved" when used as nouns to be respectful of the people we study.

[3] Linking the historical records of women remains difficult, limiting us to follow only the paternal line of ancestry. However, we assess the degree to which intermarriage between descendants of the Enslaved and the Free may affect our estimates. In sum, with plausible levels of intermarriage, the true Free-Enslaved gaps in 1940 are likely 1.5 as large as our baseline estimates.

also allows us to identify the place where a family was freed. Our second strategy does not require record linking and instead uses last names to determine how likely a family is to descend from the Enslaved (Ager et al., 2021). Specifically, we leverage changes in the distribution of last names in the census from 1860 to 1870—before and after the inclusion of the Enslaved—to assign a probability of having been enslaved until the Civil War to each last name.[4]

Using this new data, we first document the degree to which the socioeconomic status of Black families continues to depend on how their ancestors experienced slavery and Jim Crow. Specifically, we quantify the gaps in education, income, and wealth between descendants of the Free and the Enslaved between 1870 and 2000. These gaps capture differences that arise from the Enslaved both being enslaved longer (on average around 50 years) and being more likely to be freed in the Deep South—the epicenter of Jim Crow. We use full-count census records to track outcomes until 1940, the last year in which such data is available. Thereafter, we link the 1940 census to administrative records that contain a family's residential address in 2000 to proxy for the economic conditions in which they live today.[5]

Our results show that a family's socioeconomic status today continues to drastically vary by when and where they were freed. Descendants of the Enslaved have considerably lower education, income, and wealth than descendants of the Free. Throughout the first 70 years after slavery, descendants of the Enslaved remained twice as likely to be illiterate than descendants of the Free. Our results show that both in 1940 (individual level) and in 2000 (neighborhood level), the Free-Enslaved gap in various outcomes has magnitudes equal to around 40 percent of the Black-white gap.

We decompose these within-race disparities into the part that reflects differences in how families experienced slavery and the part that reflects differences in how they experienced Jim Crow. To do so, we document how the Free-Enslaved gap conditional on being freed in the same location. This conditional gap reflects the any persisting disadvantage due to being enslaved longer among families who, because they lived in the same location, faced the same Jim Crow institutions after slavery. The remainder of the Free-Enslaved gap results from the effect of having been freed in different places.

Conditional on being freed in the same location, the large initial socioeconomic gaps in 1870 between families freed at different times vanish by 1940. In 1870, Black families

[4]For example, the last name "Freedman" did not exist before the Civil War, but was chosen by many families who were newly freed in 1865. In contrast, the last name "Du Bois" became ten times less frequent in the census after the formerly Enslaved were included in 1870, suggesting that families with that name were likely free before the Civil War.

[5]These administrative death records contain a deceased person's last neighborhood of residence, which we can link to socioeconomic outcomes at the neighborhood level. The administrative data contain the neighborhood (nine-digit ZIP code) of a person at death, and they capture the near-universe of deaths between 1988 and 2005. The records linked to our 1940 census sample are almost evenly divided into deaths before and after 2000. We pool those observations and refer to it as the 2000 sample.

who gained freedom in 1865 fared substantially worse in terms of education, income, and wealth compared to families who had gained freedom earlier lived in the same location but . Over the decades, this conditional gap narrowed and vanished by 1940. This pattern implies that the observed gap between descendants of the Free and the Enslaved in 1940 is not driven by how long their ancestors were enslaved, but rather by exposure to Jim Crow institutions determined by geographic location.

A back-of-the-envelope calculation shows that due to low geographic mobility, the effect of being freed in a location is a very good proxy for the effect of living in that location after slavery. With Black families freed in the Deep South faring far worse that families freed elsewhere, it may seem puzzling that less than 15 percent of them left the slave states by 1940. The geographic mobility of Black families was especially low in the Deep South due to institutional factors—such as Jim Crow laws that limited labor mobility (Roback, 1984)—and economic factors—such as high migration costs (Carrington et al., 1996). Even if a Black family migrated to the North, opportunities there often eroded quickly (Akbar et al., 2020; Derenoncourt, 2022).

To further document how a Black family was affected by oppression under Jim Crow in the long-run, we next document how Black economic progress after slavery differed across space. Specifically, we assess how each county or state affected the socioeconomic status of Black families who were freed there in 1865. We leverage variation in the location of freedom using the birthplace of families' enslaved ancestors. Conditional on the place where their ancestors were freed, we trace the socioeconomic status of families irrespective of whether or where they move over time.

We find vast geographic differences in the long-run economic progress that Black families made after slavery depending on where they were freed. Both in 1940 (individual level) and in 2000 (neighborhood level), Black families freed in states further south tend to have a lower socioeconomic status than families freed in states further north. Families freed in the Deep South stand out with particularly low socioeconomic status. For example, they attained a full year less in education in 1940 than those freed elsewhere in the South—with the average Black person attending school for six years. The results at the county level mirror those at the state level and families across counties of the same state often made very similar amount of progress until 1940. This pattern points to the importance of factors that vary sharply across state borders—such as Jim Crow institutions.

We rule out the possibility that the geography of Black families' economic progress is driven by race-neutral factors, such as urbanization or local soil quality by comparing how Black and white families fared across locations. We find that the effect of being freed in a specific county within the South is uncorrelated with the average outcomes of white people with roots to the same county ($\rho = 0.071$). In contrast, the correlation with the average outcomes of free Black Americans with roots to the same county is high ($\rho = 0.837$).

We next further isolate Jim Crow's effect from other environmental factors that may have affected the geography of Black economic progress after slavery. We employ a border discontinuity design in comparing the economic progress of families freed across (but in close proximity to) the borders of the Deep South—the epicenter of Jim Crow (Woodward, 1955).

We find that the long-run economic progress of Black families freed just inside states with a strong Jim Crow regime was substantially lower than that of families freed just outside of those states. Across the Deep South borders, the gap amounted to 0.5 years in education by 1940. Thus, only the border differences in Jim Crow institutions account for half of the disparity (1 year in education by 1940) between families freed in the Deep South and families freed elsewhere. These differences only arose across parts of the Deep South border where the intensity of Jim Crow institutions changed substantially. We confirm that before Jim Crow, no systematic differences across those borders existed in either location characteristics—such as farm values, share of the population that was Black, or population density—or in the socioeconomic status of newly freed Black families—such as their literacy or occupational income scores. Lastly, the socioeconomic status of white families, in contrast to Black families, was not affected by which side of the Deep South border their ancestors lived on in 1870.

Next, we show that restricting the access to educational resources was a key channel through which Jim Crow slowed Black economic progress after slavery. We document that states where Black families made the least economic progress were those which implemented the most severe Jim Crow laws pertaining to schooling (but not necessarily other domains). We also show that a county's fraction of Black children in school between 1870 and 1940 is the single most robust predictor of the county's effect on Black economic progress, among a large variety of other county-specific factors, such as cotton suitability, demographic composition, or the number of lynchings. Schooling continues to stand out even if we only consider within-state variation and control for all other observable county characteristics. The high correlation of Black families' long-run economic progress and their access to schools is consistent with education being both a critical determinant of intergenerational mobility and a main target of Jim Crow.

Finally, we provide causal evidence for both access to schools being a critical determinant for Black families' long-run economic progress and a main channel of Jim Crow being the restriction of such access. To do so, we leverage quasi-random variation in whether a Black child benefited from one of 5,000 Rosenwald schools that were built throughout South between 1914 and 1931 (Aaronson and Mazumder, 2011). We assess how the access to schools affected Black children across states that differed in the intensity of Jim Crow. We also study whether the Rosenwald school program had intergenerational effects that persist until today by linking the children of benefiting individuals from the 1940 census to their administrative death records in 2000.

We find that the Rosenwald school program had persistent positive effects on the economic progress of Black families, especially in the most oppressive states. A Black child able to attend a Rosenwald school attained 0.3 years more in education than a child not able to attend. In the Deep South, the effect was more than twice as large. This finding suggests that it was not a lack of incentive to accumulate human capital that slowed Black economic progress under Jim Crow, but rather the lack of resources to do so. In addition, attending a Rosenwald school not only increased the education of the benefiting individuals, but also improved economic conditions of their children today.

In sum, our results show that the economic progress a Black family made after slavery critically depended on the location in which they were freed. The majority of families who were enslaved until the Civil War were freed in the Deep South. After slavery ended, the Deep South was the region in which state governments implemented the most severe forms of Jim Crow institutions, designed to limit the economic thriving of Black families. With its critical importance for intergenerational mobility, the denial of equal access to education through Jim Crow was a key factor that made the Deep South so detrimental to Black economic progress. While Jim Crow ended in the 1960s, its effects persist until today. Our results suggests that had it not been for Jim Crow institutions, Black economic progress after slavery would have been substantially faster.

We have argued that variation in families' location of freedom is plausibly exogenous, one may be concerned that the location in which a family was freed is partly determined by their innate characteristics. Our main identifying assumption is that an enslaved person's birthplace is exogenous to future generations' potential socioeconomic outcomes. This assumption is supported by historical evidence. The Enslaved had no freedom of movement before 1865, which limited the potential for selection into location. A remaining identification concern may be that the system of slavery itself induced a degree of selection into location of the enslaved population. Such selection was likely limited. Around half of forced migration of enslaved people occurred through slaveholder migration, which was generally non-selective as slave masters tended to move all their enslaved people with them (Fogel and Engerman, 1974; Pritchett, 2001; Tadman, 2008; Pritchett, 2019).[6] The domestic slave trade accounts for the rest of forced migration. Historical evidence suggests that the only locations that show evidence of selective slave trade were the sugar cultivation areas—accounting for about 6 percent of the rural enslaved population (Tadman, 1977, 1979). By the nature of the work required, enslaved Americans there tended to be stronger and more likely to be male (Phillips, 1918). While having able-bodied men to clear the new highly-fertile land is the kind of selection that slave owners cared about, it is not the kind of selection that would influence human capital 100 years later. Contrary to the most plausible scenario for positive selection, we find

[6]In principle, selection could also arise through differences in the slaveholders who choose to migrate. However, for selection to arise, the slaveholder's decision would need to be correlated with the potential outcomes of their enslaved people—a scenario we cannot rule out but deem unlikely.

that families freed in the Louisiana sugar areas achieved lower socioeconomic status by 1940 than families freed in other areas.

This paper contributes to our understanding of how modern racial disparities are affected by historical racial oppression. Regions—both within and outside of the U.S.—that relied on slave labor continue to have lower and more unequally distributed incomes (Nunn, 2008), lower upward mobility (Berger, 2018), larger racial disparities (Bertocchi and Dimico, 2014), and higher levels of racial resentment against Black Americans (Acharya et al., 2018).[7] Empirical estimates of how slavery and other forms of historical racial oppression affected *individual families* are scarce. Sacerdote (2005) uses census cross-sections which include the birthplace of a person's parents to show that in the early 1900s, Black Americans whose parents were likely freed during the Civil War continued to have lower socioeconomic status than those whose parents were likely freed earlier. Historical record linkage allows us to study this persistence over many generations. In contrast to prior work, our method also identifies Black families who gained freedom within the South before the Civil War (half of all free Black families). As a consequence, we can exploit rich variation in both when and where a family was freed, allowing us to study different drivers of persistence.

This paper also contributes to the historical literature on Black economic progress after slavery. Margo (1991) argues that beyond market forces such as the supply of and demand for educated Black workers, one of the main reasons why Black children did not catch up with the education of white children after slavery was their parents' poverty and illiteracy. Our results show that the "intergeneratinonal drag" of slavery was quantitatively important for around three generations. We extend his model of Black economic progress to encompass the racially oppressive institutions after slavery that we show to be decisive in the long-run. The dependence of Black economic progress on institutional factors is consistent with the seminal work of Du Bois (1935), Woodward (1955), and Ransom and Sutch (2001) who highlight that when and where their environment allowed for it, Black families did make rapid progress—such as in the Reconstruction era (1865–1877). Consistent with that, the Rosenwald program being particularly effective in the most oppressive states suggests that it was not a lack of demand for education among Black children there, but rather a lack of access to education that slowed their human capital accumulation (see also Aaronson and Mazumder, 2011).

This paper also contributes to our understanding of geographic disparities in intergenerational mobility. Upward mobility has been lower in the South than in any other region, both historically (Olivetti and Paserman, 2015) and in recent decades (Chetty et al., 2014). Consistent with this evidence, our results suggest that Black economic progress af-

[7] Those locations also have lower productivity (Mitchener and McLean, 2003). Fujiwara et al. (2019) find that in Brazil, a location's historical reliance on slave labor caused weaker institutions and higher inequality until today.

ter slavery was slowest in the South. Chetty and Hendren (2018) analyze children whose families move across places to show that locations *caused* low upward mobility in the South rather than being a result of selection. Our empirical strategy relies on the immobility of a specific population rather than their mobility. Our results confirm the importance of exposure to location-specific factors for the intergenerational accumulation of human and physical capital. Indeed, a state's capacity to generate upward mobility is highly persistent: our estimates have a strong correlation $(0.617, p = 0.000)$ with the causal effect that states have on intergenerational mobility of low-income groups in recent decades (Chetty and Hendren, 2018). The high persistence of a location's capacity to generate upward mobility is consistent with the persistence of cultural, economic, and institutional factors that set boundaries to economic opportunity.

Last, this paper adds to the evidence of whether and how historical institutions can affect economic outcomes in the long-run. Acemoglu et al. (2002); Dell (2010); Donaldson (2018); Dell and Olken (2019) show that regions can be lastingly transformed by temporary policies and institutions. In this paper, we develop innovative methods to study the impact of institutions on *individual families*, rather than regions, and apply them in the context of U.S. historical racial oppression. Such individual-level evidence allows for the mobility of families (attenuating regional differences) and can generate novel insights into the mechanisms that drive various forms of persistence. For example, in this paper we leverage cohort-specific variation in how likely a family was to benefit from the Rosenwald school program to show that Jim Crow affected Black families not by reducing their demand for education, but by limiting their access to education.[8]

2. Historical Context

In this section, we provide some historical context for the evolution of racially oppressive institutions in the U.S.—from slavery and Jim Crow until today.

2.1 Free Black Americans before 1865

In 1860, just before the Civil War (1861–1865) that led to the abolition of slavery, 4 million enslaved and 0.4 million free Black people lived in America. Enslaved people existed on American soil from the country's colonial origins in the 16th century. The roots of the free Black population trace back to 1619, when the first 20 Black people were purchased by settlers in Virginia. Little is known about their fate, but it is likely that at least some of them were treated as servants who had to work for a fixed term and gained freedom

[8]This finding is consistent with evidence from development economics, showing that the provision of educational resources is most effective in areas where prior to the intervention, children were most deprived of opportunities to access such resources (e.g., see Duflo, 2001; Glewwe and Kremer, 2006).

after (Frazier, 1949). Around 1660, both law and practice had changed, implying that virtually all Black individuals who arrived in the colonies were enslaved for life. From 1662 onwards, the law also mandated that a child would inherit their legal (i.e., free or enslaved) status from their mother regardless of race.

For those enslaved, the Revolutionary War (1775–1783) provided a road to freedom. Responding to a need for troops and laborers, the British governor promised freedom to all enslaved willing and able to serve the British. It is estimated that up to 100,000 enslaved people ran away from plantations to do so (Schama, 2006). After the war, many of them remained in the U.S. as free persons. As a result, the free Black population in some states increased dramatically.

The Revolutionary War also brought with it a spirit of egalitarianism, challenging the institution of slavery in some regions. In the North, the abolitionist movement spread quickly after the war. While only few Black people lived free of slavery before the Revolutionary War, most Northern states adopted gradual Emancipation laws after the war. New Jersey was the last Northern state to do so in 1804. By 1860, of the 226,000 Black Americans in the North, all but 64 were free.

In the South, the path to freedom was narrow, especially in the Lower South. All Southern states except North Carolina allowed masters to free ("manumit") their slaves by 1790, but the practice was employed to different degrees across regions.In the Upper South, the first wave of manumissions occurred between 1783 and 1793, the first decade after the Revolutionary War. Motivated by anti-slavery beliefs, the vast majority of manumitters freed all their slaves at once. However, manumission gradually became more selective and turned into a reward system designed to uphold slavery (Wolf, 2006). By 1860, 0.2 million of the 1.8 million Black Americans in the Upper South were free. The Lower South did not see a similar manumission wave after the war as manumissions there were usually limited to masters' "illicit offspring, special favorites, or least productive slaves" (Berlin, 1974). The free Black population of the Lower South mostly originated from refugees who fled from the French colony Saint-Domingue (now Haiti) and the purchase of Louisiana from France that had a sizable free Black population. By 1860, 40,000 of the 2.5 million Black Americans in the Lower South were free.

The legal and socioeconomic status of the Free varied greatly across locations and over time before 1865. In most states, free Black Americans were deprived of the right to vote and to hold political office. But, importantly, their property rights were legally protected and in most cases respected. With the limited freedom they enjoyed, some of the Free could accumulate modest wealth and social status. Most of them, however, lived in poverty "under conditions barely distinguishable from those of the mass of slaves" (Berlin, 1974). Their economic status varied greatly across the country and, perhaps surprisingly, tended to be better further South (Berlin, 1976). In the North, the Free were concentrated in cities where they suffered from competition with and hostility from white

laborers (Frazier, 1949). Most of the Free in the South lived in rural areas, working as farmhands and casual laborers (Berlin, 1974). Those who lived in the Southern cities were more prosperous than their rural counterparts (Jackson, 1942; Frazier, 1949).

2.2 Freedom of All Black Americans after 1865

After 1865, the formerly enslaved were emancipated, giving all Black Americans the same legal status. The average free Black American family had already been free for likely around 50 years. For the first 12 years after the Civil War—the Reconstruction era (1865–1877)—the Union Army occupied the South. This era allowed Black Americans to experience unprecedented socioeconomic progress. New schools and universities were built to educate Black Americans throughout the South. Black men participated politically, casting their votes in high numbers and serving in public office (Logan, 2020). While the period was exceptional for Black Americans as a whole, it was often the free Black Americans, not the formerly enslaved, who became political leaders or rose to high socioeconomic status in other professions (Fitchett, 1947).

Throughout the Reconstruction era, Black economic and political progress was met with violent opposition from white Southerners (Foner, 1963). In 1877, the Union troops left the South, abandoning the project of Reconstruction. What followed was the disenfranchisement of Black people through informal and legal means, leading to massive declines in Black political participation (Kousser, 1974; Naidu, 2012). Many free Black Americans lost their higher social status and left the South (Woodson, 1918).

Those who remained in the South after Reconstruction faced increasing oppression through the rise of Jim Crow (1877–1964). Jim Crow laws governed almost every aspect of Black life. Schools, work places, public transport, medical facilities, and parks were racially segregated. Black suffrage was limited through poll taxes, literacy tests, and other means. Enticement, contract enforcement, and emigrant-agent laws prevented Black workers from seeking economic opportunities with new employers or in states outside the South. Vagrancy laws criminalized the unemployment of Black people. In addition to institutionalized segregation, various informal means of excluding Black Americans spread through the South.

From 1910 to 1940, many Black Americans started to leave the (Upper) South in the first wave of the Great Migration. While the Civil Rights Movement successfully fought oppression starting in the mid-1950s, the Great Migration continued until the end of the movement in the late 1960s. By then, six million Black Americans had left the South. Even after the achievements of the 1960s, old forms of racial oppression have persisted and new forms—such as mass incarceration—have arisen (Western, 2006; Alexander, 2010; Darity et al., 2016). Progress in narrowing racial gaps in socioeconomic status have largely stalled since the 1960s (Bayer and Charles, 2018; Derenoncourt et al., 2022).

3. Data and a New Method to Identify Where and When a Family was Freed from Slavery

We construct family histories for all Black Americans in the census between 1850 to 2000. We develop a new method to identify (1) where a Black family was freed from slavery and (2) whether or not they had already been free before the Civil War (1861–1865).

3.1 Data

We use full-count census data for all available decades between 1850 and 1940 (Ruggles et al., 2020) and link observations across adjacent and non-adjacent decades using the automated linking methodology provided by Abramitzky et al. (2020). A person is linked from one census to another if their name, year of birth, and state of birth match and, importantly, if the match is *unique* conditional on race. Using these links and the relationship of individuals within households, we can create family histories that allow us to study their social, economic, and geographic mobility across generations. Our main outcomes include education, income, and wealth. Over time, the census data tend to provide increasingly rich information on those outcomes. We study the evolution of outcomes between 1870 and 1940, before focusing particular attention on the rich information included in the 1940 census.

To extend our analysis to the 21st century, we link the 1940 census to administrative death records between 1988 and 2005 (Goldstein et al., 2021). These records cover the near-universe of deaths among American citizens and contain the nine-digit ZIP code of the decedent's residence at their time of death. These codes are highly granular indicators of location, which refer to a "segment or one side of a street" (USPS, 2021).[9] Through this geographic granularity, we can obtain rich information on the socioeconomic characteristics of a person's neighborhood. We use National Historical Geographic Information System (NHGIS) data on the distribution of education, income, and wealth by race within these areas to extend our analysis to the present day. We also examine broader measures of well-being such as health, mortality, residential segregation, evictions, and fatal police encounters. Last, we build a database of historical location characteristics including agriculture, demographics, lynchings, segregation, railroad networks, and transportation costs (see Appendices A.5–A.7 for a detailed description of all variables).

[9] We map nine-digit ZIP codes into statistical areas, such as census blocks, which are small and designed to have socioeconomically homogeneous populations (Census Bureau, 2017). In Appendix A.8, we lay out the procedure to link nine-digit ZIP codes to statistical areas.

3.2 Sample

For our main analysis, we focus on Black men aged 20 to 54 years. We also limit our sample to individuals who can be linked to their ancestors in 1880 or earlier, for two reasons. First, to identify a family who gained freedom before 1865, it must be linkable to their ancestors in 1850 or 1860. Restricting the sample to Black Americans linkable to 1880 or earlier minimizes the bias that may result from comparing families who can be linked back in time easily (e.g., because they have unique names) with those who cannot. Second, this restriction excludes families who immigrated to the U.S. after 1880.

For 1940, the latest year available, our sample of Black prime-age men consists of 9,325 descendants of the Free and 155,813 descendants of the Enslaved. We achieve an average linking rate of 12 percent between adjacent census decades and 20 percent to any future decade. From the 1940 census, we can link 1,591 male descendants of the Free and 21,059 male descendants of the Enslaved to their ZIP code in the death records.

3.3 Identifying Descendants of the Free and the Enslaved

The key empirical challenge we overcome is to identify when and where a Black family was freed from slavery—i.e., how long they were enslaved and how intensively they were exposed to Jim Crow after slavery. We develop a new method to accurately distinguish two groups: descendants of the Enslaved—freed in 1865, mainly in the Deep South—and descendants of the Free—freed before the Civil War, mainly in the Upper South and North. To identify descendants of the Free, we make use of the circumstance that before 1865, the enslaved population was not recorded by the census. Any Black person in the 1850 or 1860 census was free.[10]

3.3.1 Main Classification Method: Linking Historical Census Records

Our preferred method to identify descendants of free and enslaved Black Americans is based on automated census-linking methods (Abramitzky et al., 2019). We classify individuals born before 1865 as free Black Americans if they were recorded by the 1850 or 1860 census or if they were born in a state that had already abolished slavery.[11] We classify individuals born before 1865 as (formerly) enslaved if they were both born in a slave state and not recorded in the 1850 or 1860 census. The Free-Enslaved status of Black men is then carried forward to their descendants. Appendix A.3 describes our classification algorithm in greater detail.

Our classification method has three advantages over existing work. First, it allows us

[10] These are the only pre-1865 census decades with individual-level data.

[11] The sample of the Free increases by one third after including families who we cannot link back to 1850 or 1860 but who we know to be born free based on their state and year of birth.

to study the long-run effects of differential exposure to the institution of slavery. Previous research typically relied on birthplaces to identify descendants of the Free and the Enslaved. With the census only providing this information for a person and their mother, the intergenerational effects could not be determined beyond the second generation.[12] Second, it correctly identifies the large number of free Black Americans in the South; indeed, 50 percent of the Free lived in the South before 1865. By using birthplaces as a proxy for their Free-Enslaved status, all free Black Americans with roots in the South were misclassified as descendants of the Enslaved. Third, and most importantly, it allows us to analyze differences between descendants of the Free and the Enslaved conditional on where their ancestors lived around Emancipation. We find that this geographic component is crucial to understanding the key channel—post-slavery institutions— driving the persistent disadvantage faced by descendants of the Enslaved.

Intermarriage between descendants of the Free and the Enslaved, in contrast, may attenuate our estimated Free-Enslaved gap towards zero. Because women tended to change their last names upon marriage, our Free-Enslaved classification only takes into account male ancestry. An advantage of considering only male ancestry is that it avoids bias through selective marriage. A disadvantage is that we do not have an intensive-margin in a person's Free-Enslaved status. In Appendix F, we use spouses' state of birth (and their mothers' state of birth) to show that while intermarriage was rare in absolute numbers due to the differences in the two groups' geographic concentration, it likely has a sizable impact on our estimates. This is because the population of the Free is relatively small, so that even low absolute occurrences of intermarriage have a relatively large impact on the group. We show that by 1940, the true Free-Enslaved gaps are around one-and-a-half times as large as our original estimates.

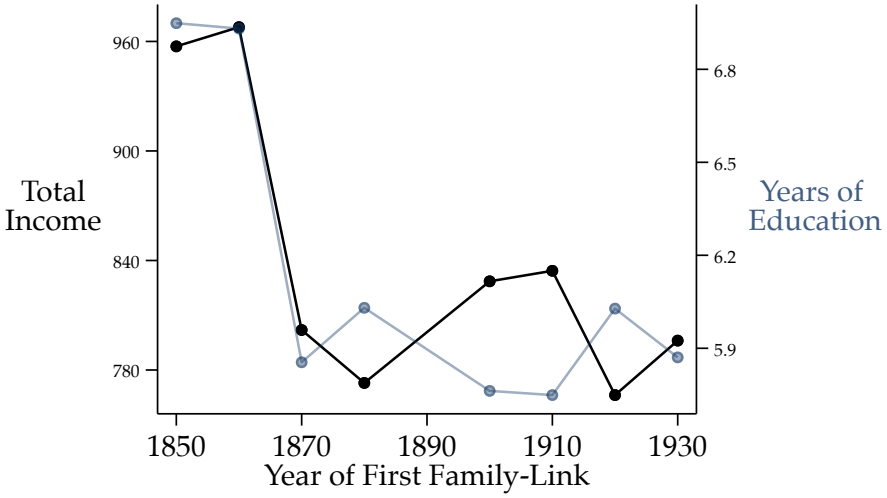
Potentially mis-classifying descendants of the Free as descendants of the Enslaved due to imperfect matching rates has a negligible effect on our results. In principle, if an individual cannot be linked back to the 1850 or 1860 census, this could either mean that they were enslaved or they could not be matched by the linking algorithm. Hence, in the South, we mis-classify some descendants of the Free. However, census records show that only 6.2 percent of the Southern Black population were free in 1860, many of whom we correctly classify as such. As a result, we compare a group that was free with almost certainty in 1860 with a group of which at least 94 percent were enslaved. Thus, the potential for attenuation bias is minimal.

[12] This extension also lets us use the rich information on education, income, and wealth in the 1940 census.

3.3.2 Potential Linking Bias

In constructing our sample, we rely on linking individuals across census records. One may be concerned that linking procedures introduce mechanical differences in descendants of the Free and the Enslaved. The most plausible concern is that a person's socioeconomic status depends on how many generations or decades they can be linked backward in time. For example, children of single mothers can typically not be linked to their grandparents in earlier censuses because existing methods exclude women.

FIGURE 1: Average Outcomes in 1940



Notes: This figure shows the average outcomes of Black Americans in 1940 by the earliest year to which we can link them back to one of their ancestors. The black line (left y-axis) shows the total predicted income; the blue line (right y-axis) shows the years of education. The lines suggest no trend in outcomes outside of the break from 1860 to 1870. See Data Appendix A for details on the sample and data.

To examine the quantitative importance of this concern, we group Black Americans in 1940 by the earliest decade in which we can link them back to one of their ancestors, and plot their average outcomes by group (see Figure 1). In 1870, the formerly Enslaved are included in the census for the first time. Consistent with that change in sample composition, we observe a large drop in average income and education of people who can be linked to ancestors in 1870 but not 1860 or 1850. Aside from this drop, there are no trends in income or education, suggesting that individuals who can be linked back further do not have mechanically higher socioeconomic status. The evidence suggests that the Free-Enslaved gap estimated here is not affected by such a mechanical bias. To err on the side of caution, we limit our sample to individuals who can be linked back to 1880 or earlier throughout this paper (for both descendants of the Free and of the Enslaved). We show that our sample is highly representative of both the Free and the formerly Enslaved by comparing our sample of the Free to Black Americans in the 1860 census (exclusively free Black Americans) and our sample of the formerly Enslaved to the Black population in the 1870 census excluding those we know to be free Black Americans (see Table A.1).

3.3.3 Alternative Classification Method: Distribution of Last Names

We develop a second strategy to identify descendants of the Free and the Enslaved based solely on last names. We use the change in the distribution over last names before 1865, when the census included only free Black Americans, to after 1865, when the it included all Black Americans. This approach allows us to use the full (rather than only the linked) sample of Black Americans in the census. The two approaches yield Free-Enslaved classifications that are highly correlated (see Appendix Figure A.3).

While some last names were common both among the Free and the formerly Enslaved, others were characteristic of one group. For example, the last name "Du Bois" was relatively frequent among Black Americans in the 1860 census. However, with the inclusion of the formerly Enslaved in the census of 1870, Du Bois became 90 percent less frequent—an indication that having this last name meant a person likely descended from the Free. In contrast, the last name "Freedman" is non-existent in the 1860 census but appears in the 1870 census after some formerly enslaved people chose it as their new last name. Thus, Black families called Freedman likely descend from the Enslaved.

This alternative classification method trades off accuracy in favor of coverage. Based on this probabilistic measure, we expect the resulting Free-Enslaved gaps to be subject to attenuation bias, making them appear smaller than they are in truth. In contrast to our main method, however, this classification can be applied to the entire population of Black Americans rather than being limited to those who we can link back in time. In addition, this classification also allows us to identify descendants of the Enslaved in non-census data that include last names. We are currently working with one of the largest credit score provider in the U.S. to extend our estimates of the Free-Enslaved gap until 2022.

4. Socioeconomic Gaps between Descendants of the Free and the Enslaved Since 1870

In this section, we document the gaps in education, income, and wealth between descendants of the Free and the Enslaved from 1870 to 2000.

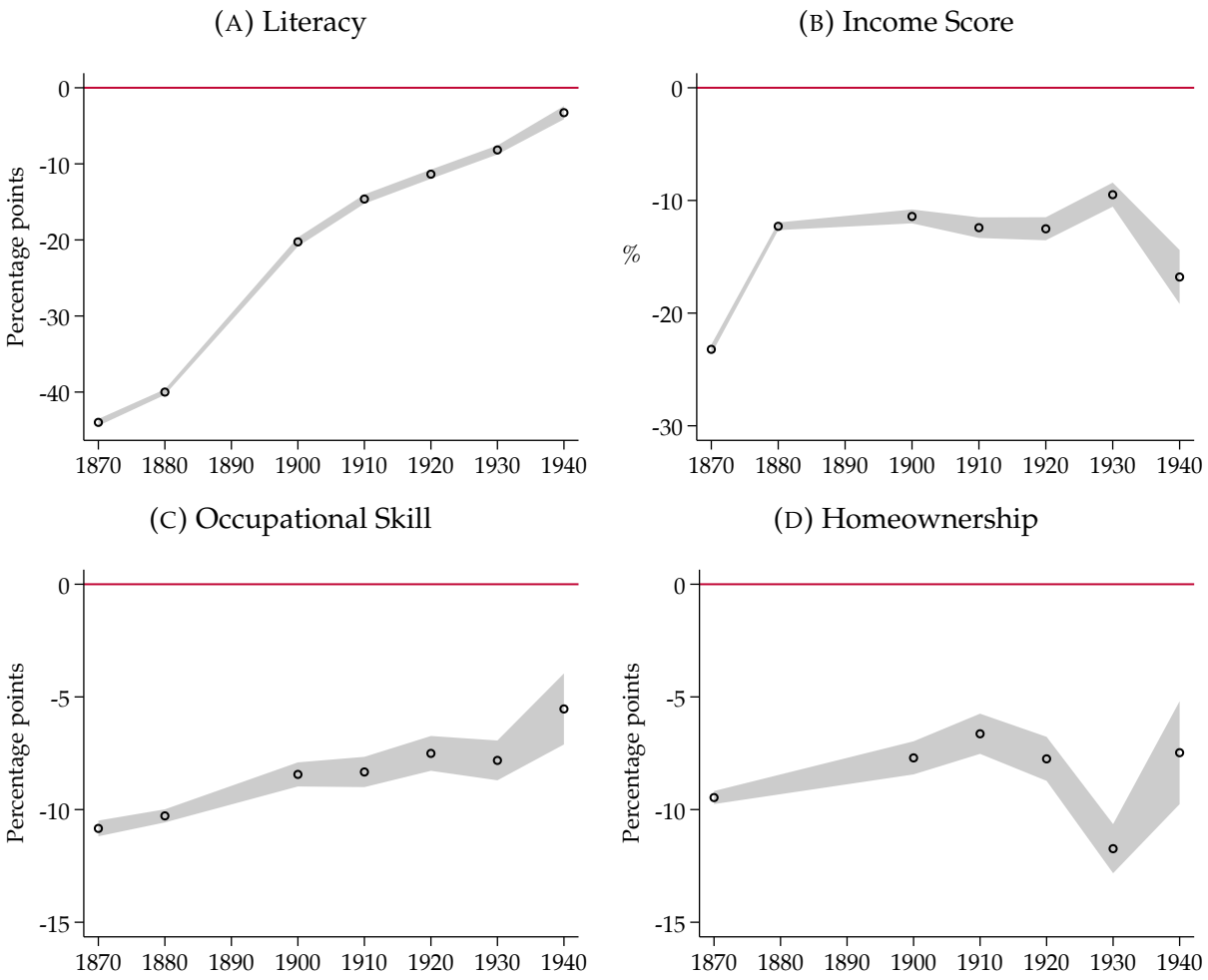
4.1 Persistence of the Free-Enslaved Gap until 1940

We first estimate the Free-Enslaved gap (β_t) in socioeconomic outcomes ($y_{i,t}$) separately for each decade from 1870 to 1940:

$$y_{i,t} = \alpha_t + \beta_t s_i + \gamma' X_{i,t} + \varepsilon_{i,t}, \quad (1)$$

where s_i is equal to one if person i is classified as a descendant of the Enslaved and is zero otherwise. $X_{i,t}$ is a vector of controls that includes a quadratic term of age in our baseline specification. We cluster standard errors at the 1870 ancestor level.

FIGURE 2: Free-Enslaved Gap (1870-1940)



Notes: This figure shows the gaps in literacy rates, income scores (OCCSCORE), occupation skill levels (HISCLASS), and homeownership rates among prime-age (20-54) male descendants of free vs. enslaved Black Americans in each census decade. We restrict the sample to observations linked to ancestors in 1850, 1860, 1870, or 1880. We control for a quadratic function in age and include 95 percent confidence bands clustered at the family level. As literacy is reported only until 1930, we use educational attainment beyond second grade as a proxy in 1940. Figure 5 shows the Free-Enslaved gap conditional on one's pre-1865 ancestor being born in the same state. See Data Appendix A for details on the sample and data.

The socioeconomic differences between descendants of the Free and the Enslaved are large and persistent. In 1870, the formerly Enslaved were 2 times (over 40 percentage points) more likely to be illiterate than free Black Americans (see Figure 2). By 1940, the gap was still 1.8 times (5 percentage points). A similarly strong persistence can be documented for income and wealth. Descendants of the Enslaved work in less skill-intensive occupations than descendants of the Free throughout 1870 to 1940. Consistent with this skill gap, descendants of the Enslaved earn substantially lower incomes and are significantly less likely to own their home.

The Free-Enslaved gap narrowed far more slowly over this period than comparable gaps among white Americans. To benchmark the speed of convergence, we estimate socioeconomic gaps between white families whose ancestors had no measurable physical or human capital in 1870 and all other white families between 1870 and 1940 (see Appendix Figure I.20). In only 30 years, the gap in literacy rapidly shrunk from over 90 percentage points (twice the Free-Enslaved gap in 1870) to less than 10 (half of the Free-Enslaved gap in 1900). The homeownership gap for the two groups of white Americans was similar to the respective Free-Enslaved gap in 1870 but closed by 1900—while the Free-Enslaved changed very little until then. A higher persistence among Black families is consistent with their low levels of intergenerational mobility throughout modern history (Collins and Wanamaker, 2022; Chetty et al., 2020).

The rich information on education, income, and wealth provided by the 1940 census allows us to get a detailed picture the Free-Enslaved gap 75 years after slavery ended. Our results in Table 1 confirm that even in 1940, descendants of the Enslaved are substantially less educated, earn lower incomes, and have accumulated less wealth than descendants of free Black Americans. The gap in education amounts to 1.6 years—more than one-quarter of the average years of education among Black men in 1940.[13] The probability that a descendant of the Enslaved earned a high school or college degree by 1940 was only half compared to descendants of the Free; for graduate degrees it was only one-third (see Appendix Table J.11).

Consistent with the educational gap, the wage incomes of the Enslaved are $146 (USD) lower, 38 percent of average incomes among Black men.[14] In terms of wealth, descendants of the Enslaved are 7.7 percent less likely to own a home, and conditional on homeownership, their houses are worth $713 less, on average—half of what the average house owned by Black Americans was worth.[15]

The alternative measure based on the distribution of last names confirms that large Free-Enslaved gaps exist in 1940. As expected, the results based on this probabilistic classification yields more attenuated gaps. The implied gaps correspond to 9 percent of the average education among Black prime-age men, 5 percent for income, 11 percent in homeownership rates, and 61 percent in house values (see Appendix Table J.12).

We next conduct an array of placebo exercises that give additional support to our empirical strategy (see Appendix Tables I.6 and I.5). Computing the Free-Enslaved gap based on 1875 (rather than 1865) as the placebo year of Emancipation, a small gap of less than 1 percent in education emerges (compared to 25 percent in our baseline). When

[13] This gap is at the lower end of the range that Sacerdote (2005) estimates based on comparing Southern- and Northern-born Black men older than 54.

[14] Appendix Table J.11 compares the Free-Enslaved gap across different measures of income. The relative gap in (imputed) *total* income is similar to the relative gap in years of education: 25 percent.

[15] In Appendix J.3, we document the Free-Enslaved gap in intergenerational mobility that has contributed to this persistence.

TABLE 1: Free-Enslaved Gap (1940)

	Education (Years) Mean: 5.99	Wage Income (USD) Mean: 381.20	Homeownership (%) Mean: 29.25	House Value (USD) Mean: 1,371.95
Ancestor Enslaved	**-1.59*****	**-145.92*****	**-7.24*****	**-694.69*****
	(0.05)	(6.13)	(0.62)	(65.85)
Controls (age, age^2)	Y	Y	Y	Y
Adjusted R^2	0.04	0.05	0.01	0.01
Observations	163,549	154,463	164,357	46,971
Ancestor Free	*9,078*	*8,551*	*9,070*	*3,227*

Notes: This table shows the gap in years of education, wage income, homeownership, and house value among prime-age (20–54) male descendants of free vs. enslaved Black Americans in 1940. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

computing the Free-Enslaved gap for white (rather than Black) Americans yields no economically (and rarely statistically) significant estimates. Adding various placebo measures as controls to our baseline specification does not affect our conclusions about the persistence of the Free-Enslaved gap.

While our estimates are consistent with prior evidence, our estimation strategy allows for a broad reinterpretation of the results. The literacy gaps we estimate are comparable to those estimated by Sacerdote (2005). For 1880, he finds that Black people born in the South before 1865 are 47 percentage points less likely to be literate than those born in the North. As expected, our estimated Free-Enslaved gap of 40 percentage points in 1880 is somewhat smaller because our classification can correctly identify free Black Americans in the South. The Southern Free have lower literacy rates than free Black Americans in the North. Similarly, Sacerdote finds gaps of around 16 percentage points for descendants in 1920, whereas we find gaps of around 11 percentage points.

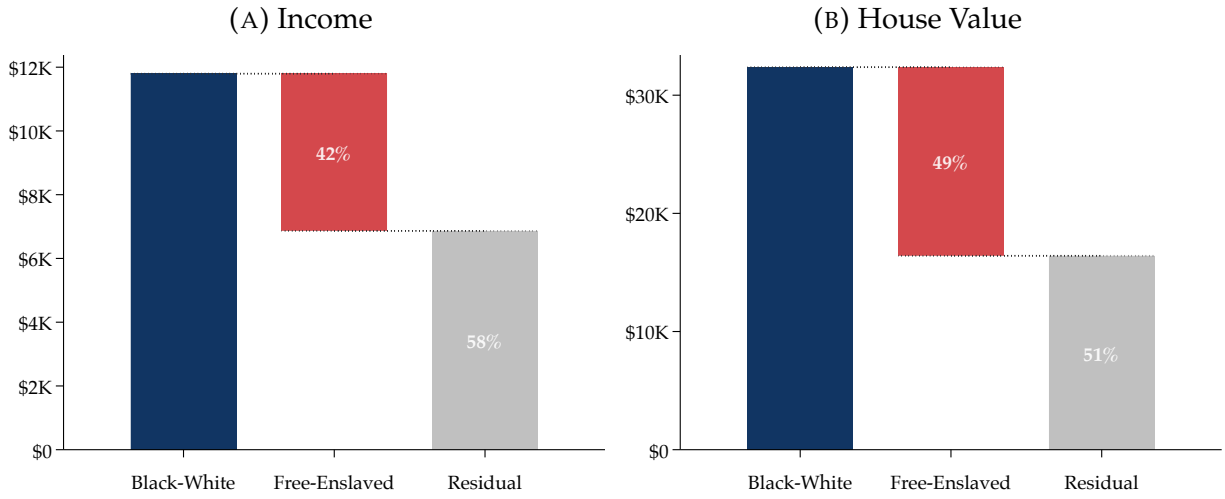
4.2 The Free-Enslaved Gap in the Early 21st Century

While our main results show that descendants of the Enslaved did not catch up with descendants of the Free until 1940, it is reasonable to suspect that the narrowing accelerated over the course of the Civil Rights Movement (1954–1968). Indeed, existing evidence suggests that the social mobility of Black Americans temporarily increased around 1970 (Clark, 2014).[16] To extend our results to more recent decades, we link Black families from the 1940 census to administrative death records around 2000. Those records include the nine-digit ZIP code of a person's residence at the time of death from which we derive

[16] Clark (2014) uses the last names of "underclass" Black Americans to study their representation among licensed physicians and attorneys, two "elite" occupations.

neighborhood-level information on the distribution of income, wealth, and education. Our sample primarily covers the 1910–1940 birth cohorts.

FIGURE 3: Decomposition of the Black-White Gap (2000) at Census Tract Level



Notes: This figure decomposes the Black-white gap among men in 2000 in income (Panel A) and house values (Panel B) into the component accounted for by the Free-Enslaved gap and a residual. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. For this decomposition, we only use variation across census tracts. House values are measured conditional on ownership. Each person is assigned the respective race- and gender-specific median across the census tract in which they lived at the time of death. Appendix Figure J.25 decomposes the gaps in income and house value based on the individual-level data of the 1940 census. See Data Appendix A for details on the sample and data.

In 2000, descendants of the Enslaved reside in neighborhoods that have substantially lower education, income, and wealth than those of the Free's descendants. Our estimates indicate that descendants of the Enslaved are 3.9 percentage points less likely to hold a high school degree and 2.6 percentage points less likely to hold a college degree (see Appendix Table J.9). Their expected income is lower by $5,100—around 17 percent of the Black median income. Conditional on owning their home, their houses are worth around $17,500 less—around 19 percent of the median house value among Black Americans.[17] Because those estimates ignore within-neighborhood differences, they should be seen as a lower bound for the Free-Enslaved gap.

In 2000, the Free-Enslaved gap has magnitudes equal to over 40 percent of the modern Black-white gap in income and wealth (see Figure 3) and one-third of the Black-white gap in college attainment (see Appendix Figure J.24). Importantly, even free Black Americans were almost certainly harmed by slavery. Most had been enslaved at some point in their family's history, and their descendants faced racial discrimination during and after slavery as a result of the institution. It is merely the differences in how they experienced slavery and Jim Crow that accounts for over 40 percent of the Black-white gap. Next, we turn to the drivers of this persistence.

[17] We also show that descendants of the Enslaved live in neighborhoods associated with one year less in life expectancy and with a 16 percent higher chance of a fatal police encounter (see Appendix Table J.16).

5. Drivers of the Free-Enslaved Gap

This section proposes a conceptual framework to understand why the socioeconomic status of Black families continues to differ by their ancestors' historical experience of racially oppressive institutions. We propose a simple model of intergenerational mobility that incorporates the importance of both when and where a family was freed while allowing for different forms of selection.

5.1 A Simple Model of Black Economic Progress after Slavery

We propose a simple model to guide the interpretation of the Free-Enslaved gap and to analyze the channels that drive its persistence. The framework incorporates intergenerational mobility, the effects of exposure to location-specific factors, (potentially selective) migration, and the effect of delayed freedom. We discuss the following questions: Is the persisting disadvantage faced by descendants of the Enslaved a causal effect of slavery or Jim Crow? Could selection into being a free Black American drive a large part of the gap? What role did the differential exposure to location-specific factors among the Enslaved and the Free have in shaping the gap?

Let $y_{i,t}$ denote the income—or any other outcome of interest—for family i at time t. For simplicity, let there be two time periods, $t \in \{0, 1\}$; the model is easily extendable to more time periods. We think of $t = 0$ as reflecting 1865, the year of Emancipation, and $t = 1$ as reflecting 1940, the last census year that we can link people to. We model $y_{i,t}$ to be determined by

$$y_{i,t} = \alpha_{i,t} + \gamma^t_{l(i,t)} + \rho_{l(i,t-1)} y_{i,t-1} + \varepsilon_{i,1} \tag{2}$$

such that it depends on five factors: a factor capturing innate ability $\alpha_{i,t}$ with c.d.f. $F(\cdot)$, the family's previous income $y_{i,t-1}$, their location $l(i,t) \in \mathcal{L}$, an indicator for whether the family was enslaved until Emancipation s_i, and a random error term $\varepsilon_{i,t}$ that satisfies $\mathbb{E}[\varepsilon_{i,t} \mid s_i, \alpha_{i,t}, l(i,t)] = 0$. Last, we define γ^t_l as the effect of being exposed to factors specific to location l at time t. The mobility parameter ρ_l can vary across locations. We model $y_{i,0}$ (the starting condition) as

$$y_{i,0} = \alpha_{i,0} + \gamma^0_{l(i,0)} - \delta s_i + \varepsilon_{i,0}. \tag{3}$$

That is, in 1865, the outcomes depend on ability, location, and whether a person had been free before the War. The parameter $\delta \geq 0$ may capture any direct advantage that free Black Americans had relative to the enslaved population, such as access to education and the right to earn an income or accumulate wealth during slavery. At time $t = 1$, the

outcomes then become

$$y_{i,1} = \left(\lambda + \rho_{l(i,0)}\right)\alpha_{i,0} + \rho_{l(i,0)}\gamma^0_{l(i,0)} + \gamma^1_{l(i,1)} - s_i\rho_{l(i,0)}\delta + \rho_{l(i,0)}\varepsilon_{i,0} + \varepsilon_{i,1}, \tag{4}$$

where we have imposed that $\alpha_{i,1} = \lambda\alpha_{i,0}$. That is, we allow for intergenerational transmission of ability over multiple generations. We see that outcomes over time are determined by ability of the initial generation through direct inheritance of ability (λ) and through intergenerational advantage derived from ability in previous generations ($\rho_{l(i,0)}$). The current location ($\gamma^1_{l(i,1)}$) shifts the level of a person's income. Through intergenerational transmission, incomes are also affected by (1) how previous generations were affected by where they lived ($\gamma^0_{l(i,0)}$), (2) whether their ancestors were enslaved until the Civil War (δ), and (3) their ancestors' idiosyncratic income shocks ($\varepsilon_{i,0}$).

We think of the effect of being enslaved until the Civil War as the expected difference between the two groups holding constant any confounding factors (i.e., ability). That is, we define the average "treatment" effect (ATE) at $t = 1$ as

$$\begin{aligned} ATE \equiv \int \left(\mathbb{E}[y_{i,1} \mid s_i = 1, \alpha_{i,0}] - \mathbb{E}[y_{i,1} \mid s_i = 0, \alpha_{i,0}]\right) dF(a_{i,0}) = \\ = \int \mathbb{E}\left[\rho_{l(i,0)}\left(\gamma^0_{l(i,0)} - \delta\right) + \gamma^1_{l(i,1)} \middle| s_i = 1, \alpha_{i,0}\right] dF(a_{i,0}) - \\ \int \mathbb{E}\left[\rho_{l(i,0)}\gamma^0_{l(i,0)} + \gamma^1_{l(i,1)} \middle| s_i = 0, \alpha_{i,0}\right] dF(a_{i,0}). \end{aligned} \tag{5}$$

Importantly, in our definition, the effect of descending from an enslaved person includes not just the effect of delayed freedom but also any potential effect operating through differential exposure to location-specific factors. Combining (3), (4), and (5), the observed Free-Enslaved gap is equal to

$$\mathbb{E}[y_{i,1} \mid s_i = 1] - \mathbb{E}[y_{i,1} \mid s_i = 0] = ATE - B, \tag{6}$$

where the (negative of) the selection bias B, arising from (1) potential selection into being free, (2) potential selection into location by (descendants of) the Free, and (3) potential selection into location by (descendants of) the Enslaved (formally defined in Equation (15) in Appendix B.1). Naturally, if there was no selection into the status of being a free Black American, the observed Free-Enslaved gap becomes an unbiased estimate of the ATE in (5):

$$\mathbb{E}[y_{i,1} \mid s_i = 1] - \mathbb{E}[y_{i,1} \mid s_i = 0] = ATE \quad \text{if } \alpha_{i,0} \perp\!\!\!\perp s_i. \tag{7}$$

We discuss the historical evidence on the plausibility of this assumption in Appendix B.1. In summary, whether someone was enslaved until 1865 was in many cases a result of factors largely exogenous to their "ability." For instance, half of the free population became free through general emancipation in the North. In the South, freedom was attained in different ways, and some were more likely to be selective than others.

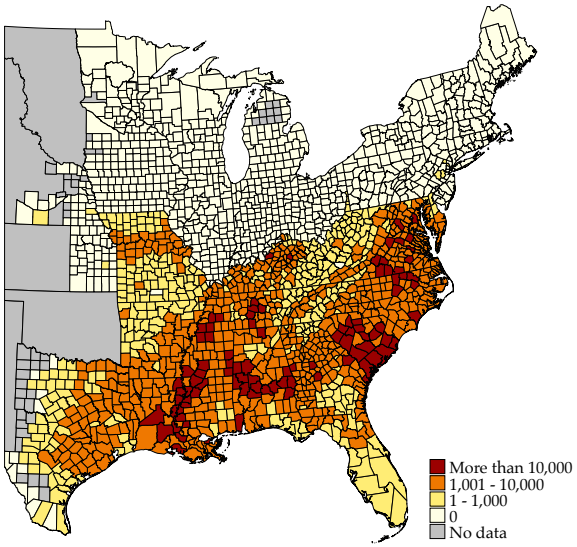
To assess the relative importance of being enslaved longer, but facing the same set of Jim Crow institutions after, we next examine the Free-Enslaved gap conditional on the location of the ancestor living in 1870. This analysis also defines a tight upper bound on the bias B in Equation (6).

5.2 The Free-Enslaved Gap Conditional on Ancestor Location

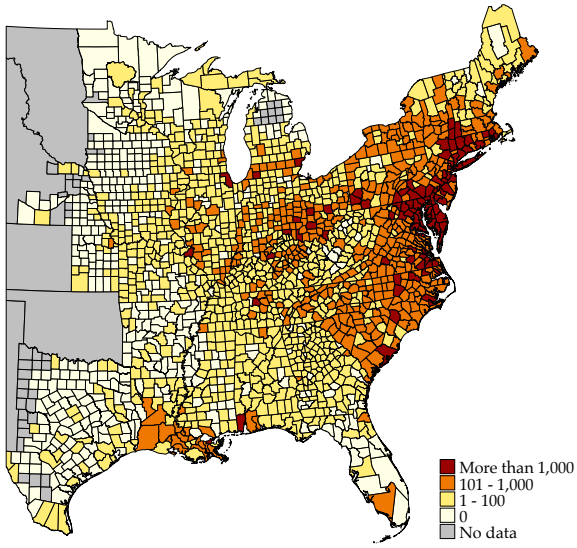
While most of the enslaved population was freed in the Lower South, most of the Free lived in the North or Upper South by the end of slavery (see Figure 4). Before and after slavery ended, these regions differed in terms of their economies, institutions, and politics. We explore to what extent the Free-Enslaved gap can be accounted for by the different environments that the formerly Enslaved and the Free lived in.

FIGURE 4: Population by County in 1860

(A) Number of Enslaved



(B) Number of Free



Notes: This figure shows the population sizes of enslaved Black Americans (Panel A) and free Black Americans (Panel B) in the 1860 census. The maps are truncated to omit the western half of the country, which at the time was only sparsely populated. The full maps are shown in Appendix Figure D.9.

To analyze the importance of locations, we control for the location that i's ancestors lived in around Emancipation as documented by the 1870 census. Formally, we add state fixed effects (η) to our initial specifications for outcomes (1870–2000),

$$y_{i,t} = \alpha + \beta_1 s_i + \eta_{l(i,1870)} + \gamma' X_{i,t} + \varepsilon_{i,t}, \tag{8}$$

where $l(i, 1870)$ is the state of birth of i's ancestors in 1870. The estimate of β_1 effectively compares 1940 or 2000 outcomes between descendants of the Free and the Enslaved conditional on their families originating from the same state.

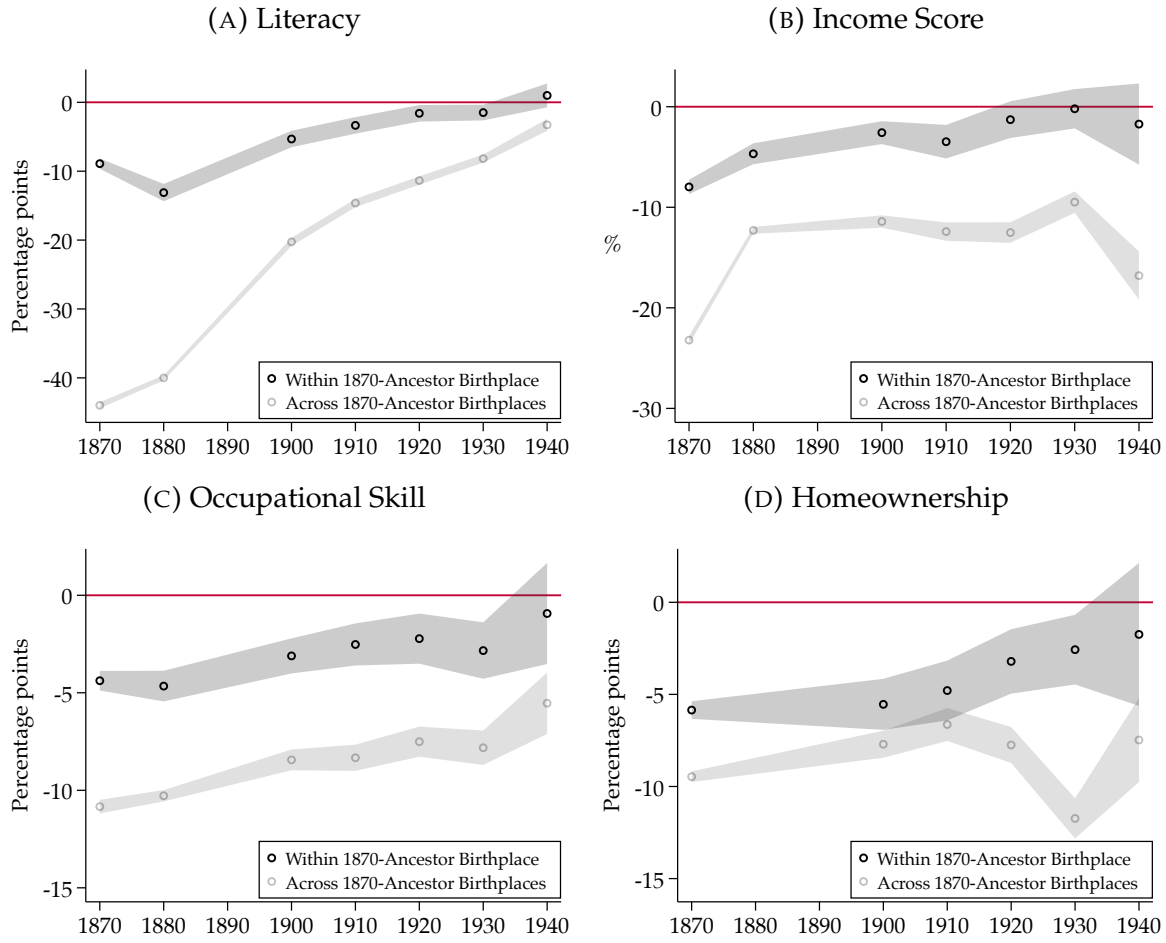
Within-state of origin gaps between the Free and the Enslaved disappeared by 2000 and most had already disappeared by 1940. The 1940 gaps in literacy, income scores, skill, and homeownership shrink to virtually zero after conditioning on 1870 origin (see Figure 5). Only the gap in years of education remains statistically and economically significant until 1940: descendants of the Free received 0.4 years more education than descendants of the Enslaved from the same state of origin (see Appendix Table J.14). The 1940 difference in wage income shrinks from $137 to $20 after controlling for state of origin. We detect no conditional differences in homeownership or house values in 1940. By 2000, the Free-Enslaved gap conditional on 1870 ancestor state suggests full convergence (see Appendix Table J.15). The finding of Free-Enslaved convergence within state of origin in 1940 or 2000 is not specific to any region, but hold across regions (see Appendix Figure J.23 for gaps by region from 1870 to 1940 and Appendix Figure J.22 for gaps by state in 1940)

Our estimates suggest that whether or not a family was enslaved until the Civil War (1861–1865) or became free earlier had a large impact on its socioeconomic status in 1870, but the effect vanished after around three generations. The large initial effect and its vanishing is consistent with Sacerdote (2005), who argues that by 1940, the disadvantage of delayed freedom conditional on location ceased to affect descendants of the Enslaved. More broadly, this finding is consistent with the importance of geographic differences in explaining disparities in intergenerational mobility between different groups throughout modern American history (Chetty et al., 2014; Connor and Storper, 2020; Ward, 2020; Abramitzky et al., 2021; Tan, 2022).

Besides suggesting that the role for any within-location effects seems to have been modest at most, the small gaps conditional on location can also be informative for the degree of "selection bias" in the Free-Enslaved gap. As illustrated in Equations (6) and (15), the Free-Enslaved gap can differ from the average treatment effect of being enslaved until 1865 if the Free constituted a selected group. In particular, one may entertain the possibility that a person with higher unobserved ability may have had a larger probability of becoming a free Black person before 1865. Formally, the assumption that $\alpha_{i,0} \perp\!\!\!\perp s_i$ in Equation (7) may not hold, and part of the large Free-Enslaved gap could be a result of selection into being free.

The state of birth of a person's 1870 ancestor accounts for virtually the entire Free-Enslaved gap by 1940 and 2000, respectively (see Appendix Tables J.14 and J.15). With most—if not all—of the variation in the Free-Enslaved gap accounted for, there is little potential for positive selection into being a free Black American to affect the gap. This conclusion is based on two assumptions. The first is that the pre-1865 geographic distribution of the Free and the Enslaved was a result of the peculiarities of the institution of slavery and not itself driven by selection into being free. The second assumption is that the Free in the South were on average, if anything, *more* positively selected than the Free

FIGURE 5: Geography and Socioeconomic Gaps between Descendants of the Free and the Enslaved (1870-1940)



Notes: This figure shows the gaps in literacy, income, skill, and homeownership before (light) and after (dark) including fixed effects for 1870 ancestor state of birth. The comparison is made between prime-age (20-54 years) male descendants of free vs. enslaved Black Americans in each census decade. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included, minimizing bias due to the fact that the Free by definition have a link to 1850 or 1860. Both panels control for age and include 95 percent confidence bands clustered at the family level. See Data Appendix A for details on the sample and data.

in the North.[18] In Appendix B.2, we use our model of Section 5.1 to formally show that the estimates in Appendix Tables J.14 (for 1940) and J.15 (for 2000)—all small or zero—are upper bounds to the (absolute) selection bias in the Free-Enslaved gap.[19] In sum, our evidence suggests that the Free-Enslaved gap can be interpreted as a causal effect for a Black American to personally descend from the Enslaved, capturing both that the Enslaved were freed later and that they were more likely to be freed in areas where Black people faced more intensive Jim Crow institutions after slavery.

To highlight the fact that selection into freedom before the Civil War is unable to ac-

[18] See Assumptions B1 and B2 in Appendix B.2.

[19] In Appendix B.2, we formalize the set of sufficient assumptions and elaborate on their plausibility. Under these assumptions, we show formally that the ordinary least squares estimator constitutes a bound on the bias B in (6).

count for a substantial share of the Free-Enslaved gap, we also compare the Enslaved to those Free whose ancestors had not accumulated any measurable physical or human capital by the end of slavery. Even for this subsample, we find large Free-Enslaved gaps in 1940 (see Appendix Table J.13). Those differences again mostly vanish if we condition on the 1870 ancestor's birthplace.

In methodological and conceptual contrast to prior work (e.g., Sacerdote, 2005), we argue that the average treatment effect of slavery should not be thought of merely as an effect *conditional on location*. Descending from an enslaved person made a person much more likely to come from (and still live in) environments that were relatively harmful to their economic progress. The location of enslavement was directly caused by their own enslavement status, and its impact should thus be included in the treatment effect. From an econometric perspective, geographic location can then be interpreted as a *bad control* since it is a mediating variable through which slave status affected future descendants.

6. The Geography of Black Economic Progress

In this section, we document how the long-run economic progress of Black families differed by the state or county they were freed in. We focus on the formerly Enslaved to hold constant *when* families were freed.

6.1 Location of Emancipation and Its Intergenerational Effect

To identify the effect that different locations had on Black economic progress, we leverage the circumstance that the Enslaved did not enjoy freedom of movement until they were freed in 1865. The lack of free movement limits the potential for selection into locations during slavery. We estimate the effect of emancipation location c on available socioeconomic outcomes by using

$$y_{i,t} = \eta_{c(i,1870)} + \gamma' X_{i,t} + \varepsilon_{i,t} \qquad \text{if } i \text{ descendant of the Enslaved,} \tag{9}$$

where η_c is the effect of one's 1870 ancestors being born in state c while enslaved, which we use as a proxy for their location of emancipation in 1865. In our baseline specification, the vector of controls $X_{i,t}$ includes a quadratic function of age.

Given the model introduced in Section 5.1 and under the assumption that a person's enslavement location is exogenous to their unobserved ability ($\alpha_{i,0} \perp\!\!\!\perp l(i,0)$ if $s_i = 1$),

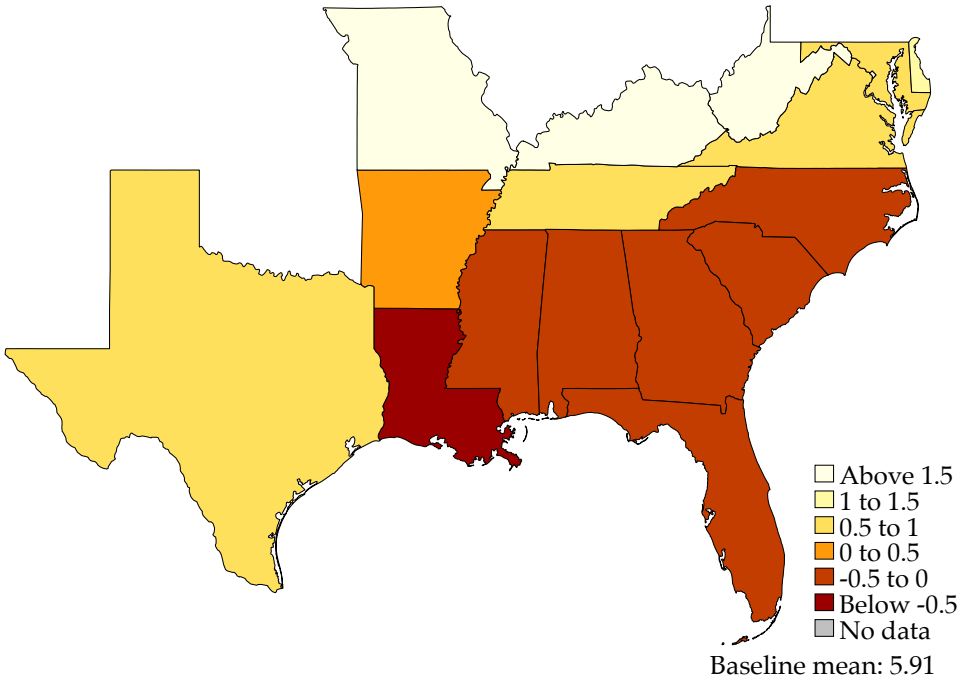
$\hat{\eta}_c - \hat{\eta}_{c'}$ is a consistent estimate of

$$\mathbb{E}[y_{i,1} \mid s_i = 1, l(i,0) = c, X_{i,t}] - \mathbb{E}[y_{i,1} \mid s_i = 1, l(i,0) = c', X_{i,t}] = \rho_{l(i,0)}\gamma^0_c + \mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 1, l(i,0) = c, X_{i,t}] - \left(\rho_{l(i,0)}\gamma^0_{c'} + \mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 1, l(i,0) = c', X_{i,t}]\right). \tag{10}$$

That is, the estimate reflects both the (inherited) effect the state of birth c had on the ancestor in 1870 (relative to state of birth c') and the expected effects of future locations given the 1870 location. One can interpret the object in (10) as an intent-to-treat (ITT) effect since the initial location is plausibly randomly assigned, but the later location (e.g., in 1940) is a result of endogenous (and potentially selective) migration decisions.

6.2 Black Economic Progress Across States

FIGURE 6: Effect of Ancestor's State of Emancipation on 1940 Years of Education



Notes: This figure shows the 1870 ancestor state of birth fixed effect (FE) estimates on 1940 years of education for descendants of the Enslaved. A state's FE is the deviation from the population-weighted average across all states (baseline mean) after controlling for a quadratic function of age. The sample includes only Black prime-age (20–54) men whose ancestors can be located in 1870. See Data Appendix A for details on the sample and data.

We find that a formerly enslaved family's state of origin has a significant impact on the outcomes of their descendants in 1940. Figure 6 shows that, generally, the states further north provided more fertile ground for Black economic progress after slavery than those further south. The state of emancipation can affect the education of descendants in 1940 by multiple years. Families freed in the Deep South stand out with particularly low

socioeconomic status. They attained a full year less in education in 1940 than those freed elsewhere in the South—with the average Black person attending school for 5.9 years.

We also consider the literacy, income, and homeownership of families in 1940 and find that the Deep South consistently provided the least amount of upward mobility for formerly enslaved families (see Appendix Figure J.26). The effect on 1940 literacy rates varies by more than 10 percentage points when comparing Louisiana (Deep South) with Maryland (Upper South). Being freed in some Deep Southern states led families to attain lower homeownership by more than 20 percentage points compared to families freed in states of the Upper South. Incomes vary by magnitudes as strong as two-thirds of the average income among Black Americans.

The negative effect of being freed in the Deep South persists until 2000 (see Appendix Figure J.27). For example, descendants of people freed in Louisiana today have average incomes that are lower by more than one-quarter of the average income among Black Americans compared to those rooted in, say, Maryland.

6.3 Black Economic Progress Across Counties

We repeat our estimation procedure at the county level by using the county of residence of a person's ancestors in 1870 instead of their 1870 ancestors' state of birth. Consistent with our state level evidence, we find that counties in the Deep South tended to be less favorable to Black economic progress than those in the rest of the South (see Figure 7). The effects of locations are spatially correlated, sometimes closely resembling the shape of states. This pattern points to the importance of state-level factors, including state-specific policies and institutions, in determining how favorable a location was to Black economic progress.

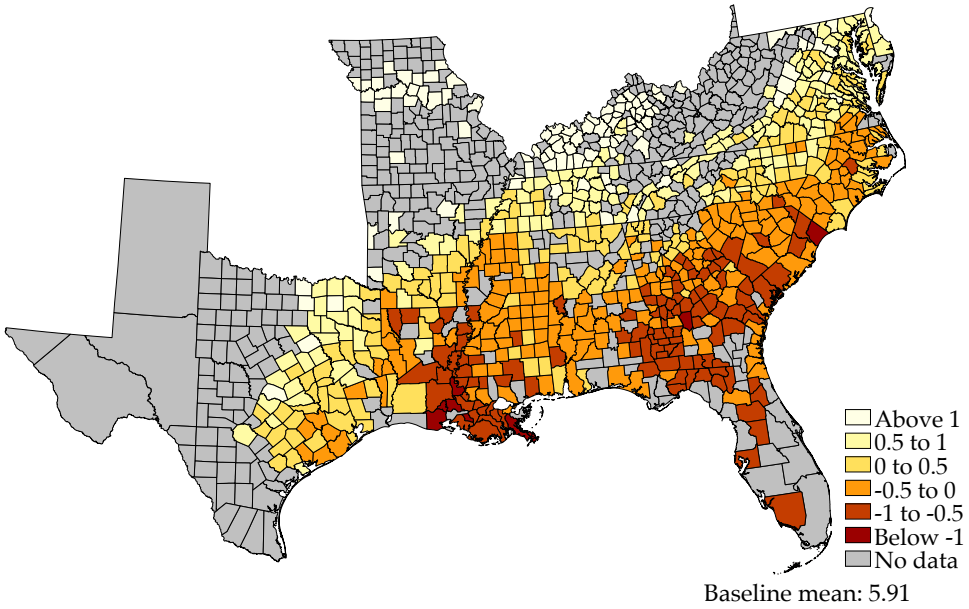
In any study that sets out to estimate a such large number of parameters, some estimates may be imprecise. We therefore reduce estimation noise by shrinking all county level estimates using a standard empirical Bayesian shrinkage method to reduce estimation noise (Armstrong et al., 2021, for details, see Appendix H).

6.4 Race and the Effect of Locations

We next examine if the effect of place is specific to race or if white Americans fared similar across locations. We compare the causal effect that each county had on the intergenerational progress of formerly enslaved families from 1870 to 1940 to the (non-causal) effect the same county had on white Americans and descendants of free Black Americans.

We find that counties that were detrimental to the economic progress of formerly enslaved families did not generally inhibit economic progress, but they inhibited *Black* economic progress specifically (see Figure 8). Specifically, Panel (A) shows that the average

FIGURE 7: Effect of Ancestor's County of Emancipation on 1940 Years of Education



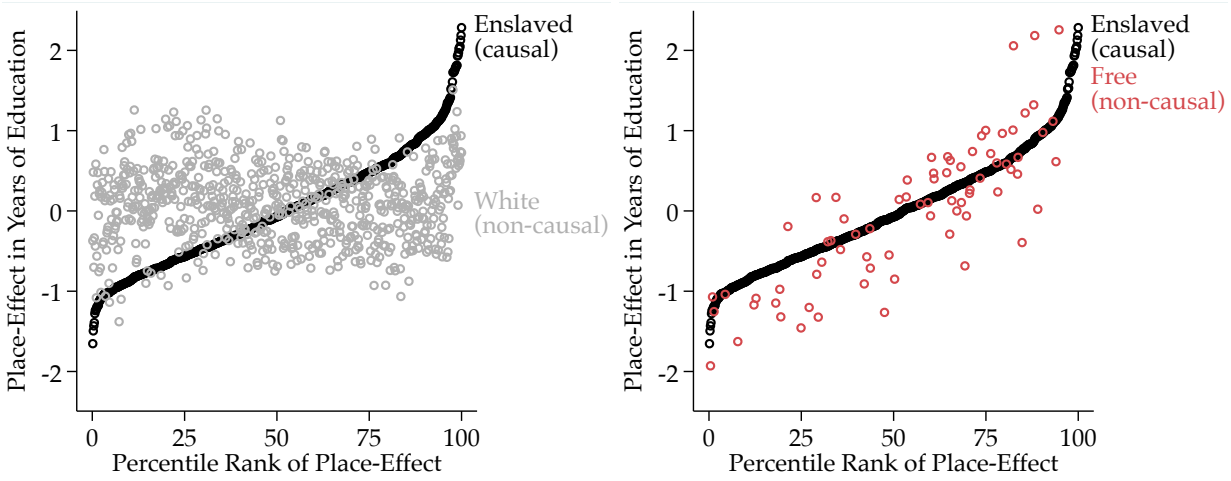
Notes: This figure shows the 1870 ancestor county fixed effect (FE) estimates on 1940 years of education for descendants of the Enslaved. A county's FE is the deviation from the population-weighted average across all counties (baseline mean) after controlling for a quadratic function of age. The sample includes only Black prime-age (20–54) men whose ancestors can be located in 1870. All county FEs based on ten observations or less are discarded. We use empirical Bayesian shrinkage as described in Appendix H. See Data Appendix A for details on the sample and data.

1940 outcomes by 1870 location for white Americans are almost entirely uncorrelated with the (causal) county effects on descendants of the Enslaved ($\rho = 0.05$). The same is true for the subsample of white Americans who descend from the poorest groups in 1870 (see Appendix Figure I.18).[20] Panel (B) shows that, in contrast, the average outcomes by ancestor location for descendants of the Free strongly correlate with the (causal) estimates for descendants of the Enslaved ($\rho = 0.84$).

The finding that a county's capacity to generate upward mobility is race specific is consistent with prior results. Jung (2021) uses counties' suitability for cotton production to instrument for local slave-to-population ratios and shows that slavery lead to depressed returns to education among Black Americans (see also Naidu, 2010) but not among white Americans. Using border discontinuities, he further shows that these disparities were the result of post-Reconstruction legislation at the state level that specifically targeted the economic mobility of Black Americans. Baker (2022) shows that states with more oppressive historical racial regimes—such as Jim Crow laws—today have higher poverty rates among Black Americans but not white Americans.

[20] This result is conditional on counties being in the South, because there, we can use plausibly exogenous variation in the locations of enslaved ancestors to estimate causal place effects. Pooling North and South, Chetty and Hendren (2018) find that the causal place effects on Black and white Americans are positively correlated. Consistent with their evidence, we find that both white Americans and descendants of free Black Americans achieve higher socioeconomic status in the North in 1940.

FIGURE 8: Ancestor Location Effects by Race



Notes: This figure compares the 1870 ancestor county fixed effects on years of education in 1940 for descendants of the Enslaved (causal) with those of white Americans and the Free (both non-causal). A county's fixed effect is the deviation from the race-specific population-weighted average across all Southern counties after controlling for a quadratic function of age. The sample includes only prime-age (20–54) men whose ancestors can be located in 1870. We use empirical Bayesian shrinkage as described in Appendix H. See Data Appendix A for details on the sample and data.

6.5 Migration and the Effect of Locations

The effect of locations that we estimate reflect how Black families fared after being freed there on average, irrespective of whether or not they migrated over time. These effects are thus affected by how likely families were to migrate, where they went conditional migrating, and how living in the different destinations (including the original location) affected Black families. Black families freed in the Deep South may have made less economic progress either because living in the Deep South harmed them or because it was harder to migrate to places with more opportunities—such as the distant Northern cities.

Indeed, migration patters differed strongly across regions. The vast majority of Black families freed in the Deep South stayed in the slave states until at least 1940, while large fractions of those freed in the Upper South left early on.[21] By 1940, the vast majority of Black families freed in the Deep South remained in one of the slave states, in some states—such as Louisiana—even over 90 percent.

We formally decompose the long-run effect of being freed in a given state into the component that is due to living there and the component that is due to migrating elsewhere. Under the assumption that the experiences that a family made in location c during slavery cease to directly affect their descendants by 1940 ($\rho_{l(i,0)}\gamma_c^0 = 0$), we can recover the average treatment effect (ATE) that living in location c had on the economic progress of Black families, γ_c^1. Equation 10 then suggests that ITT effect of initial location c, η_c, is the average over all potential future locations ATE, $\gamma_{c'}^1$, weighted by the probability of

[21] See Appendix Figure D.8.

migrating from c to c'. We invert the migration probability matrix to recover the ATE of living in each state (see Appendix G).

The decomposition suggests that the effect of being freed in a given state is a good proxy for the effect of living there after slavery.[22] Partly due to the low geographic mobility of most families before 1940, the ATE of living in each state are almost identical in most cases to the original ITT estimates of being freed in each location. For some states in the Upper South, deviations emerge but the main pattern stands: Living in the Deep South had a large negative effect on the economic progress of Black families compared to families who lived in other parts of the South.

With Black families freed in the Deep South faring so much worse that families freed elsewhere, it may seem puzzling why so few of them left the region. Both institutional and economic factors explain why the vast majority stayed for generations. First, Jim Crow directly targeted the geographic mobility of Black people (Roback, 1984): Enticement laws and contract enforcement laws limited Black workers' ability to terminate their employment contracts; Vagrancy laws essentially criminalized not having a job; Emigrant-agent laws prevented employers from seeking workers from other states. Second, moving to the North was costly, especially from the Deep South. Their geographic distance to the North limited the potential for social networks that would lower the cost of migration (Carrington et al., 1996). Even for those Black families who did migrate to the North, opportunity there often eroded (Akbar et al., 2020; Derenoncourt, 2022).

6.6 Location of Freedom and the Question of Exogeneity

In any study that aims to estimate the effect of places on intergenerational progress, the key assumption for causal interpretation is for location to be orthogonal to potential outcomes. In the literature, researchers have typically relied on "mover designs" (e.g., Chetty et al., 2016). In those studies, the effects of places is estimated from the outcomes of families who move between them. The causal interpretation derives from assumptions on the nature of their moves.

Our empirical strategy relies on the immobility of a specific population rather than their mobility. In particular, we build on the circumstance that the Enslaved did not enjoy freedom of movement before 1865, limiting the potential for selection into locations. This circumstance lends plausibility to the key identifying assumption of an enslaved person's birthplace to be orthogonal to potential outcomes of their (third-generation) descendants. The main threat to our identification assumption is the possibility of selective slave migration. Even though the Enslaved did not choose where they lived, it is possible that decisions by their owners or traders induced selection into locations of enslavement. In 1790, around 57 percent of the enslaved population lived in Maryland or Virginia. Be-

[22] See Appendix Figure G.13.

tween 1790 and 1860, there was a large shift in the geographical distribution of slavery away from the Upper South to the Lower South (see Figure D.9).

There are two ways in which this redistribution of the enslaved population could have introduced selection into location. First, it is possible that the probability of migration depended on the enslaved person's characteristics, as perceived by those who forced them to migrate. Selection into migration during the expansion of the Southern plantation economies could bias the estimated effect of the states in the Lower South relative to the Upper South. However, most enslaved people migrated with their owners. Estimates of the share of inter-regional slave migration that went through slaveholder migration vary from roughly 40 percent (Tadman, 1979) to 84 percent (Fogel and Engerman, 1974), with more recent estimates lying somewhere in between (Pritchett, 2001). Migrating slaveholders were generally non-selective in taking all their slaves with them (Fogel and Engerman, 1974; Tadman, 2008; Pritchett, 2019), leaving little room for selection into the Lower South through this type of migration. In principle, selection could also arise through differences in the slaveholders who choose to migrate. However, for selection to arise, the slaveholder's decision would need to be correlated with the potential outcomes of their enslaved people—a scenario we cannot rule out but deem unlikely.

Slave trade accounts for the remaining inter-regional slave mobility. In contrast to slaveholder migration, there is some evidence that migration through slave traders was selective: traded slaves were found to be disproportionately likely to be young adults (e.g., Pritchett, 2019) and more likely to be male (Fogel and Engerman, 1974), but some of this evidence is nuanced by Tadman (1977, 1979). Moreover, Pritchett (2001) finds that traded slaves were, conditional on age and sex, marginally taller than the average enslaved population—although this finding is contested by Steckel and Ziebarth (2016). There is no evidence that traders selected enslaved people on anything other than such basic physical characteristics. This is consistent with the dehumanization and commodification of Black people that characterized the slave trade, which "reduced people to the sum of their biological parts" (Smallwood, 2008, p. 43).

Second, conditional on migrating, the destination of an enslaved person itself may be endogenous to the potential outcomes of their descendants in 1940 or 2000. Again, given that slaveholder migration was non-selective, any such selection must have come through the slave trade.[23] Tadman (1977) examines various data sources to gauge the selectivity of the trade to different locations within the slave-importing regions of the Lower South. He finds that the only locations that show clear evidence of selection (on age and sex) are those in the sugar-producing area of Louisiana. This is consistent with the assertion of Phillips (1918) that the work on sugar plantation was generally done

[23]Given that slaveholders generally took their entire slave-holding household with them, selection through slaveholder migration could only have arisen in case slaveholders decided where to migrate based on the characteristics of the people they held enslaved (which we deem implausible).

by "able-bodied workers," which was the result "partly of selection, partly of the vigorous exertion required." However, only an estimated 6 percent of the rural enslaved population was employed in sugar production in 1850 and 96 percent of cane sugar was produced in just 20 counties.[24] In contrast to the sugar industry, the cotton and tobacco industries (accounting for around 87 percent of enslaved agricultural workers) was generally non-selective on age and sex (Tadman, 1977). There is no evidence that there was any systematic selection of enslaved people across locations other than that based on the specific crops that were being grown there. Also, as is true for selection into the slave trade, there is neither quantitative nor qualitative evidence that selection into specific locations occurred above and beyond that on age, sex, and physical appearance.

Overall, selective migration among the Enslaved may, if anything, have an attenuating effect on our estimates of place effects. We argue that possible selection on age and sex into certain areas could hardly have confounded being in those areas with the potential outcomes of one's descendants. In comparison, it seems more plausible that height and body strength could have induced some selection bias. One might hypothesize that the selection into location based on physical traits might have biased upward the estimates of states that supposedly selected positively on height and strength. However, we find that such states—those in the Lower South in general and those in the sugar region of Louisiana in particular—were particularly detrimental to Black economic progress. Moreover, such physical characteristics were co-determined by environmental influences such as nutrition, illness, the type of work, and stress (Steckel, 1979; Carson, 2008).

Lastly, for our county level evidence, we rely on a family's county of residence in 1870, around five years after Emancipation. It is possible that those who were freed in 1865 had already sorted into specific locations by 1870. Census data show that in 1870, 96 percent of children under the age of 10 lived in their state of birth (75 percent of the population lived in a household with at least one such child).[25] Therefore, it seems that across-state migration was limited. It does not, however, rule out within-state mobility across counties. Indeed, there is some evidence of within-state migration patterns in the years after Emancipation. We acknowledge that this may bias our results for certain counties in certain states (but not across states).[26] We are currently digitizing historical documents that directly record the county in which a family was freed, such as the "Virginia Census of Colored People." This data can be linked to the census to assess how mobility of Black people between 1865 and 1870 affects our county level results.

[24]The share of enslaved people employed in sugar production is taken from (DeBow, 1854, p. 94). The distribution of cane output by counties is computed from 1860 census data.

[25]We compute this statistic for households with children to avoid picking up mobility through forced migration that may have occurred long before 1865.

[26]To provide evidence that cross-state mobility between Emancipation and 1870 does not bias the results, we also compute the effects of location by the state of residence of a person's 1870 ancestor and compare it to the estimates based on state of birth.[27] Comparing Appendix Figures J.26 and J.28 shows that the effect of place on education and income are very similar when using state of birth or state of residence.

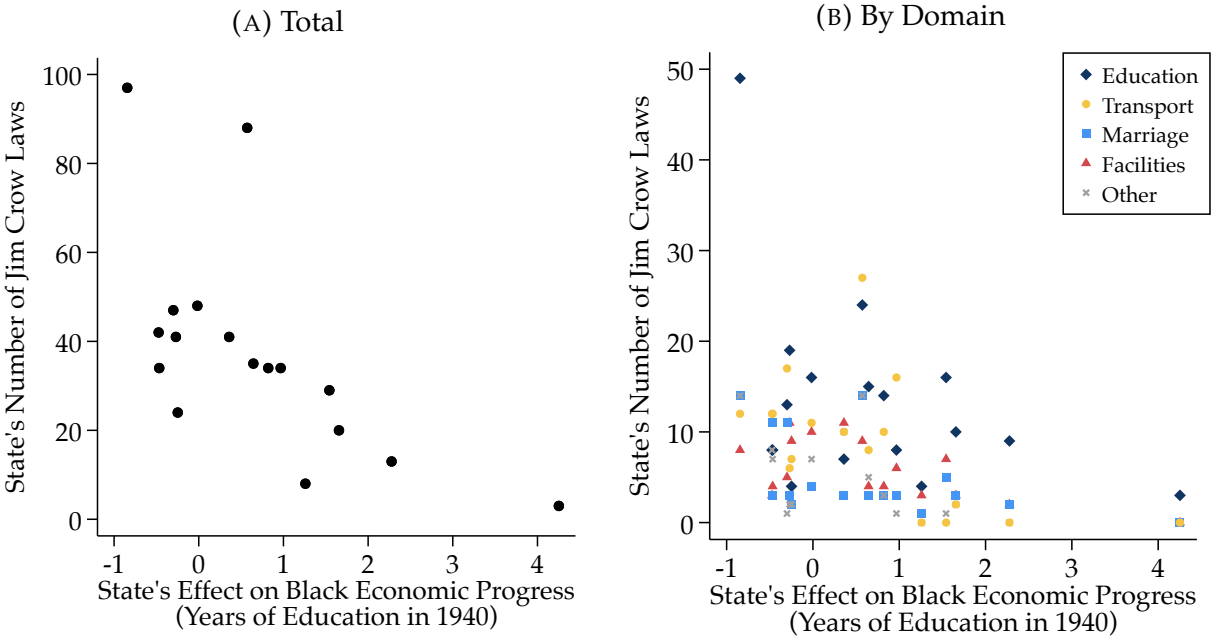
7. Jim Crow and Black Economic Progress

In this section, we study the degree to which geographic disparities in Black economic progress after slavery are attributable to Jim Crow. We employ two empirical strategies: a border discontinuity design around the Deep South and time series evidence that assesses whether institutional changes coincided with changes in Black families' socioeconomic status.

7.1 State Effects and Intensity of Jim Crow

We first document the relationship between a state's number of Jim Crow laws and the state's effect on Black economic progress after slavery. To do so, we newly digitize the laws recorded in "States' Laws on Race and Color" (Murray, 1950)—a resource aiming to document all state laws that regulated race and color in 1950. We classify each law as discriminatory, not discriminatory, or anti-discriminatory based on the legislative text and sometimes context provided by the author. We use discriminatory laws to proxy for Jim Crow laws. We also determine the domain to which each law pertains, including education, marriage, transport, facilities, and others.

FIGURE 9: Each State's Effect on 1940 Years of Education and its Number of Jim Crow Laws



Notes: This figure compares the 1870 ancestor state fixed effect estimates on years of education in 1940 for descendants of the Enslaved (causal) with each state's number of Jim Crow laws. A state's fixed effect is the deviation from the race-specific population-weighted average across all Southern counties after controlling for a quadratic function of age. The sample includes only prime-age (20–54) men whose ancestors can be located in 1870. See Data Appendix A for details on the sample and data.

The more Jim Crow laws a state had passed by 1950, the slower was the economic

progress of Black families who had been freed there (see Figure 9). Louisiana is both the state that passed the highest number of Jim Crow laws—almost 100—and the state where Black families made the least progress in terms of education by 1940. Its Jim Crow laws pertaining to education stand out as particularly high. In other states, too, education is among the domains most regulated by Jim Crow. These patterns suggest that restrictions to the education of Black children is a plausible channel through which Jim Crow limited the economic progress of Black families.

7.2 Institutional Differences Across State Borders

To isolate the role of Jim Crow era institutions (1877–1954) in shaping the geography of Black economic progress after slavery, we employ a regression discontinuity design. Specifically, we compare the socioeconomic status of families whose ancestors were freed on different sides of (but in close proximity to) state borders. While it is not only institutions that shape the causal effect of living in a given location, factors such as culture or the economy typically do not respect state boundaries. This empirical strategy thereby allows us to disentangle the effect of institutions specific to a state or region.

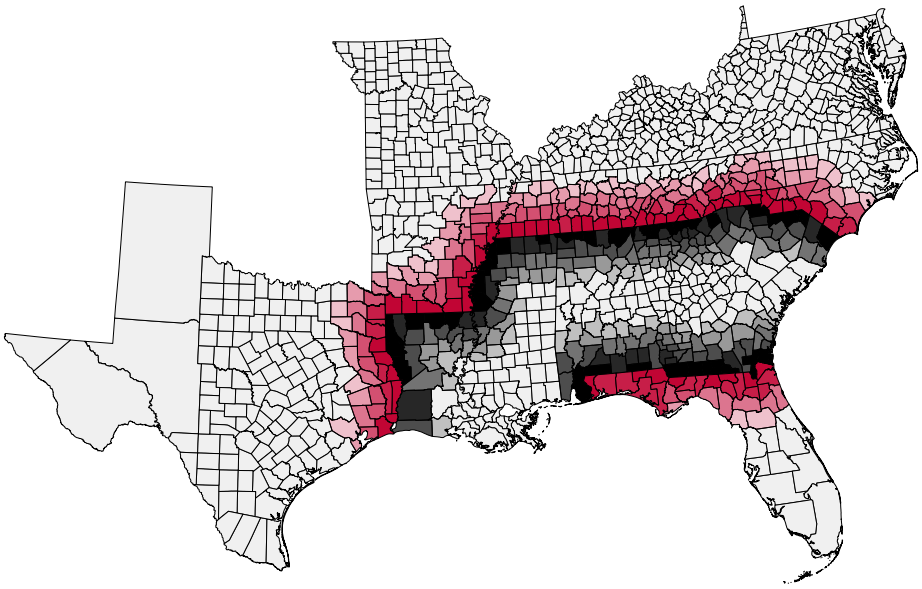
7.2.1 Deep South Border

We first employ this regression discontinuity design around the Deep South's borders (see Figure 10). While all Southern states instituted racial segregation, comparing the Deep South to other states yields a useful lower bound for the effect of institutions on Black economic progress. The Deep Southern states are frequently cited as the epicenter of de jure segregation during the Jim Crow era (Woodward, 1955). The Historical Racial Regime (HRR) score helps quantify this fact by providing a proxy for how severe institutionalized racial oppression was in a state historically (Baker, 2022). The states of the Deep South are the top-five states in terms of the HRR score.

We confirm that during slavery, location characteristics evolved smoothly across the Deep South border (see Figure 11 and Appendix Figure J.35). Specifically, while the Deep South had a relatively larger Black population, more slaves, and more cotton production, those differences did not vary discontinuously across the Deep South border. Importantly, the literacy rates of the formerly Enslaved also did not discontinuously differ across the border immediately after slavery. This started to change only as Jim Crow took hold of the (Deep) South.

The onset of Jim Crow institutions gave rise to a large discontinuous socioeconomic disadvantage among formerly enslaved families who were rooted just inside the Deep South. In 1880, the first decade of Jim Crow, literacy rates started to discontinuously drop by 9 percentage points for Black people rooted just inside the Deep South compared to

FIGURE 10: Border Discontinuity Design around the Deep South

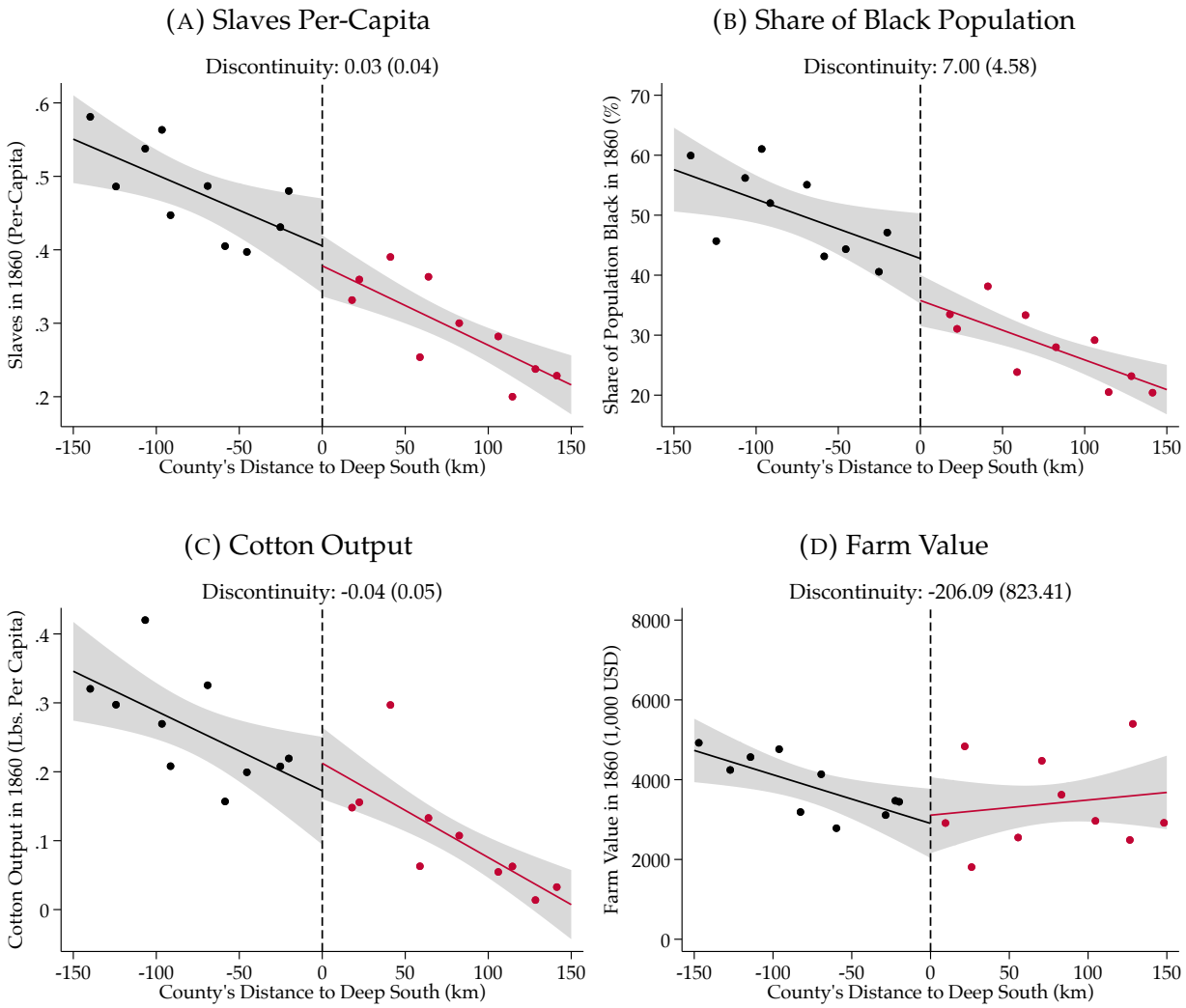


Notes: This figure shows each county's distance to the border of the Deep South. Darker shades correspond to closer proximity to the border. Counties in the Deep South are shaded in black (Alabama, Georgia, Louisiana, Mississippi, and South Carolina). Counties outside of the Deep South are shaded in red (Arkansas, Florida, North Carolina, Tennessee, and Texas). Distances are measured from a county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border.

those rooted just outside the region (see Appendix Figure J.31). By 1940, this gap was still 2.6 percentage points. While the literacy gap shrunk in absolute terms, it remained large in relative terms as 91 percent of prime-age (20–54) Black men were literate at that time. Years of education, available for the first time in the 1940 census, provide similar evidence (see Figure 12). The discontinuity suggests that a gap of 0.5 years in education (8 percent of the average among Black men) can be attributed to institutions of the Deep South. By this measure, at least half of the Deep South's uniquely negative effect (one year of education) can be accounted for by institutional differences to the rest of the South. We also find large gaps in high school and college completion, incomes, and wealth (see Appendix Figures J.32, J.33, and J.34).

We also show that the estimated discontinuities are largest in areas where racially oppressive institutions changed most dramatically across the Deep South border. Column 2 of Table 2 shows that the baseline regression discontinuity estimate of the intergenerational effect that enslavement in the Deep South had on the years of education in 1940 (as shown in Figure 12) is robust to letting the levels and trends of the years of education vary flexibly across all border state pairs. Column 3 shows that the border discontinuity effect is fully accounted for by border differences in the HRR score of the two adjacent states. A one standard deviation increase in the HRR score gap (0.97) leads to a decrease in education by 1 year. The estimates also show that there are no discontinuities across the border for states that do not differ in their HRR score.

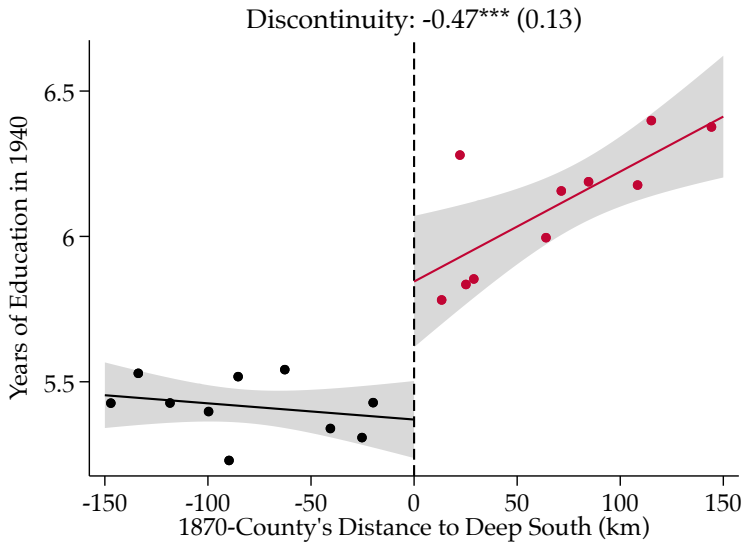
FIGURE 11: No Border Discontinuities in Location Characteristics During Slavery



Notes: This figure shows binned scatter plots of counties' characteristics in 1860 by their distance to the border of the Deep South. Lines represent the best linear fit of the individual-level data. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from a county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

There are at least two reasons why—in principle—the institutions of the Deep South may also harm white Americans. First, some Jim Crow laws are likely to harm some poor white Americans directly. For example, poll taxes aimed at disenfranchising Black voters but likely also disenfranchised many poor white voters. Second, the negative effect that institutions had on Black Americans may have spillover effects. For example, slowing Black economic progress may hamper a region's economic development more generally. Empirically, we find very little evidence that the Deep Southern institutions harmed white Americans. Appendix Figure I.19 shows that among white men, there are no border discontinuity effects (precise zeros) on literacy and years of education.

FIGURE 12: Border Discontinuities in Years of Education in 1940



Notes: This figure shows a binned scatter plots of years of education among prime-age (20-54) men in 1940 by their ancestor's distance to the border of the Deep South in 1870. Lines represent the best linear fit of the individual-level data. Shaded areas represent 95 percent level confidence bands with clustering at the county level. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from the county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

TABLE 2: Heterogeneity in Border Discontinuities

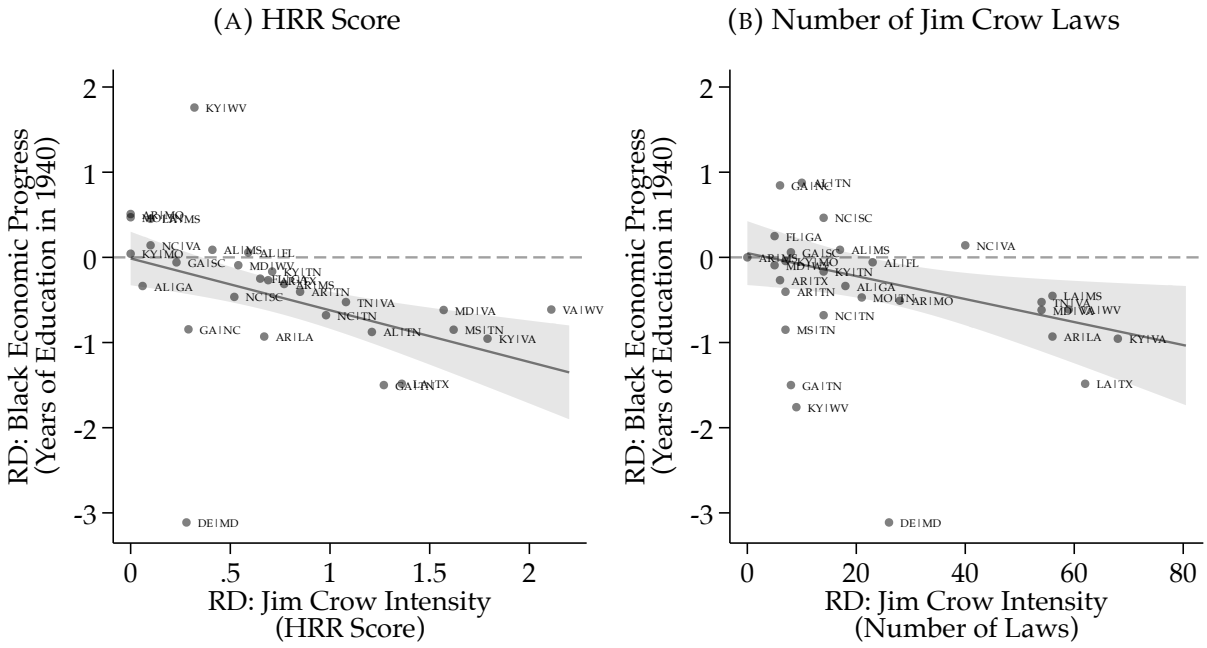
	Years of Education in 1940 Sample mean: 5.60		
Border Discontinuity	**-0.47*****	**-0.50*****	0.10
	(0.13)	(0.11)	(0.24)
Border Discontinuity × Δ HRR Score			**-0.68*****
			(0.22)
State-pair FE	N	Y	Y
Adjusted R^2	0.01	0.02	0.02
Observations	33,422	33,422	33,422

Notes: This table shows the regression discontinuity estimates around the Deep South border in years of education for prime-age (20–54) male descendants of enslaved Black Americans in 1940. Column 1 shows the estimated discontinuity show in Figure 12. Column 2 adds state pair fixed effects while allowing the slopes at each side of the border to vary flexibly across pairs. Column 3 adds an interaction between the border discontinuity and the difference in the Historical Racial Regime (HRR) score between the state pair across the border. Again, the slopes at each side of the border are allowed to vary with the HRR score difference. A high HRR score corresponds to strong racial oppression. See Data Appendix A for details on the sample and data. Standard errors are clustered at the county level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

7.2.2 All State Borders

We also employ our regression discontinuity design separately for each border between Southern states and compare the resulting discontinuities with the intensity of Jim Crow institutions across each border. We find that the larger the differences in racial oppression across two states, the larger the socioeconomic gap that opens up between families freed on different sides of the state border (see Figure 13). Families across state borders where the intensity of Jim Crow institutions does not differ reach a similar socioeconomic status. This results confirms the critical role of Jim Crow in shaping Black economic progress after slavery.

FIGURE 13: Border Discontinuities in Years of Education in 1940



Notes: This figure shows each estimated border discontinuity between two Southern states in years of education among prime-age (20-54) men in 1940 by their ancestor's distance to the state border in 1870. The linear fit is weighted by the sample size of each state pair. The shaded area represents a 95 percent level confidence band with clustering at the county level. Distances are measured from the county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

7.3 Summary: Jim Crow and Black Economic Progress After Slavery

In sum, our evidence suggests that racially oppressive institutions played a critical role in shaping the Deep South's detrimental effect on Black economic progress. The estimates are a lower bound to the importance of Jim Crow institutions for two reasons. First, all Southern states, not just those in the Deep South, segregated areas of public life by law. Our estimates isolate the *additional* effect that Deep Southern institutions had beyond Jim Crow in surrounding states of the South. Second, "laws are not an adequate index

of the extent and prevalence of segregation and discriminatory practices in the South. The practices often anticipated and sometimes exceeded the laws" (Woodward, 1955, p. 102). Our empirical strategy captures only institutions that change discontinuously at the border—such as de jure segregation at the state level—leaving some institutional features unaccounted for. The aggregate importance of institutions is almost certainly larger.

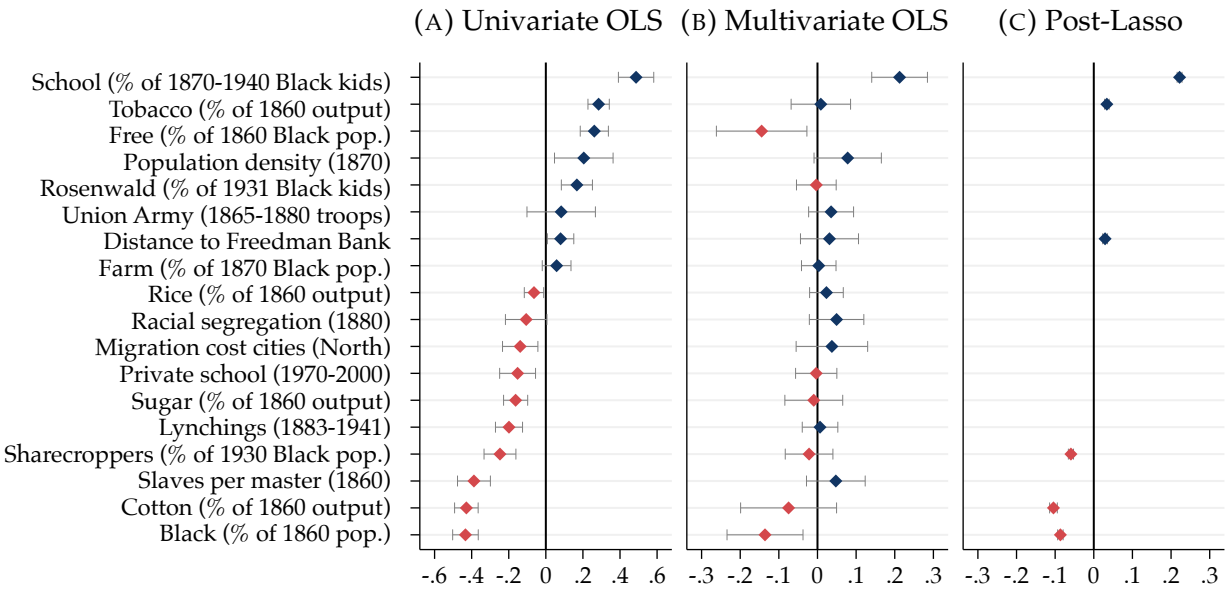
8. Education Under Jim Crow

In this section, we assess the degree to which Jim Crow limited the long-run economic progress of Black families by restricting their access to education.

8.1 Local School Attendance Predicts Black Economic Progress

We document that a county's fraction of Black children who attended school after 1865 was a key predictor for the Black economic progress families made after being freed there. First, we explore correlations between an array of location characteristics and our county fixed effects. Because some of the characteristics may be highly correlated, we also explore their predictive power in a multivariate regression that pools all of the characteristics and controls for longitude, latitude, and state fixed effects.

FIGURE 14: Correlates of Black Economic Progress by 1940

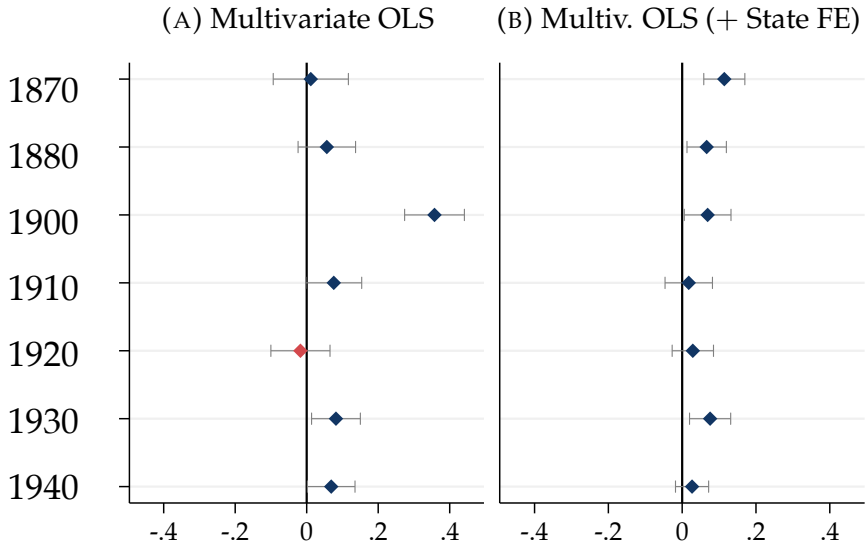


Notes: This figure shows how different location characteristics predict the 1870 ancestor county effects on 1940 years of education. Panel (A) shows the standardized coefficients from regressing the county fixed effects on each location characteristic individually. Panel (B) repeats (A) pooling all location characteristics and controlling for state FEs and a county's longitude and latitude (adjusted $R^2 = 0.63$). Panel (C) repeats (B) post-Lasso (adjusted $R^2 = 0.63$), where the penalty is chosen via the Bayesian information criterion (BIC). Coefficients are measured in standard deviations of the respective variable. Robust 95 percent confidence bars are added in black. See Data Appendix A for details on the sample and data.

Panel (A) of Figure 14 shows that counties most favorable to Black economic progress after slavery provided high rates of schooling to Black children after 1865. The population of "good" counties were also whiter, and among the Black people present there, a higher number were free. The positive correlation between our place effects and the local share of free Black Americans in 1860 is consistent with Logan (2020), who shows that the Free had positive impacts on local public good provision during Reconstruction. "Good" counties tended to be those that cultivated tobacco rather than sugar or cotton before 1865, and the slave masters there tended to own fewer slaves, on average. After slavery, fewer Black people became sharecroppers in "good" counties. Fewer lynchings also occurred in there.

Panel (B) suggests that schooling among Black children between 1870 and 1940 as a predictor for Black economic progress after slavery is robust to controlling for various other location characteristics, controlling for a quadratic function of latitude and longitude, and to considering only within-state variation. One standard deviation of the fraction of Black children in school is associated with a significant 0.2 standard deviations more years of education in 1940 when comparing otherwise identical counties. Panel (C) shows that after selecting predictors using Lasso, schooling continues to stand out as the most quantitatively important predictor of Black economic progress.

FIGURE 15: Share of Black Children in School and Black Economic Progress by 1940



Notes: This figure shows how the share of Black children attending school in each decade predicts the 1870 ancestor county effects on 1940 years of education. Panel (A) shows the standardized coefficients from regressing the county fixed effects on the pool of shares across years (adjusted $R^2 = 0.26$). Panel (B) repeats (A) controlling for state FEs (adjusted $R^2 = 0.62$). Coefficients are measured in standard deviations of the respective variable. Robust 95 percent confidence bars are added in black. See Data Appendix A for details on the sample and data.

When considering the fraction of Black children in school separately by decade (1870–1940), we find that schooling in 1900 is particularly decisive for a location's impact on

Black economic progress (see Panel A of Figure 15). This finding is consistent with the fact that only four years prior, the landmark supreme court decision "Plessy v. Ferguson" of 1896 ruled that racial segregation was constitutional, supporting the rise of the "separate but equal" doctrine. The "equal" part of the decision was rarely respected. Our finding suggests that the counties with the lowest shares of Black children in school after this decision also were the counties that were most harmful to Black economic progress after slavery. Consistent with state laws playing a major role in shaping the provision of schooling to Black children, the high impact of schooling in 1900 is completely accounted for by state fixed effects (see Panel B of Figure 15).

The intergenerational importance of education is consistent with Margo (1991), who shows that literacy and education among Black Southerners was linked to improved educational outcomes of their children, an early transition out of the farm economy, and an increased likelihood of moving to the North. Similarly, Collins and Wanamaker (2022) point to education as the key factor in shaping Black-white disparities in intergenerational mobility throughout the 20th century. Overall, the the included location characteristics account for almost two-thirds of the variation in causal place effects ($R^2 = 0.633$).

8.2 Building New Schools Increased Black Economic Progress

We leverage quasi-random variation from the Rosenwald school program to assess how schools impacted the long-run economic progress of Black families across different institutional environments. This program built 5,000 new schools for Black children throughout rural communities in South between 1914 and 1931. Aaronson and Mazumder (2011) document that there was substantial quasi-random variation in when and where Rosenwald schools were built and that counties where Rosenwald schools were built do not systematically differ from other rural counties in the South. However, our research design is robust to unobserved differences between treated and untreated counties.

Our empirical strategy follows Aaronson and Mazumder (2011) to exploit cohort- and county-specific variation in the likelihood that a Black child was able to attend a Rosenwald school. We only use variation from comparing Black children whose ancestors were freed in the same county but who differed in how likely they were to benefit from the program depending on how many of their school years (age 6–16) overlapped with a newly built Rosenwald school. In particular, we estimate

$$Y_{i,1940} = Rosenwald_{c_{1870},y} + \eta_{c_{1870}} + \eta_{s_{1870},y} + urban_{i,1870} + \varepsilon_{i,1940},$$

where $Y_{i,1940}$ is the 1940-outcome of child i born in year y, c_{1870} is the county in which i's ancestors lived in 1870, $Rosenwald_{c_{1870},y}$ is the fraction of Black children born in y who could attend a Rosenwald school in their ancestors' county c_{1870}, $\eta_{c_{1870}}$ and $\eta_{s_{1870},y}$ are

county and state-cohort fixed effects, and $urban_{i,1870}$ is a dummy for whether i's ancestors lived in an urban part of county c_{1870}.[28] Our main outcome is years of education, but we also consider literacy and proxies for income and wealth.

By leveraging the fact that our data follows families from 1850 to 2000, we make three empirical innovations. First, we use the county in which a Black child's enslaved ancestors were freed, not their own county of birth, to determine the likelihood it attended a Rosenwald school. Holding the place of treatment constant at the ancestor level addresses the concern that parents may sort into locations based on educational opportunities there.[29] Second, by following the children of potential Rosenwald school students from the 1940 census to administrative records in 2000, we analyze whether the Rosenwald schools also affected the children of those who attended. Last, we use our measures of Jim Crow intensity to explore the heterogeneity of treatment effects across different institutional environments.

It is theoretically ambiguous whether Rosenwald schools are more or less beneficial to Black children in racially oppressive places. On the one hand, the demand for human capital among Black families may be lower in oppressive areas, potentially *reducing* the impact of Rosenwald schools. For example, Jim Crow institutions may have disincentivized human capital accumulation by limiting Black children's return to education. On the other hand, oppressive institutions may directly restrict the supply of educational resources, potentially *increasing* the impact of Rosenwald schools.

We find that the provision of schools had large and persistent effects on Black families, especially in the most oppressive areas (see Table 3). A Black child who could attend a Rosenwald school had 0.31 years more in education by 1940 than a child who could not. In the Deep South, attending a Rosenwald school increased education by 0.66 years—more than twice the effect those schools had in other Southern states. Limiting the Deep South to the three most oppressive states—Louisiana, Mississippi, and South Carolina—the effect is even larger: 0.83 years. These findings suggest that Black Americans' human capital accumulation was slower in the Deep South because Jim Crow directly restricted states' supply of educational resources, not because the returns to education were low.

We confirm that only families who stayed close to where their ancestors were freed

[28] Our continuous treatment measure reflects the fraction of Black children in county c who were able to attend a Rosenwald school in the years in which child i from c was in school ages (6–16):

$$Rosenwald_{c_{1870},y} = \frac{1}{10}\sum_{t=y+6}^{y+16}\frac{RosenwaldTeachers_{c_{1870},t}\times 45}{BlackChildren_{c_{1870},t}},$$

where y is the year in which a child was born, c_{1870} is the county in which i's ancestors lived in 1870, $RosenwaldTeachers_{c_{1870},t}$ and $BlackChildren_{c_{1870},t}$ are the number of Rosenwald teachers and the number of Black children in school ages (6–16) in county c_{1870} at time t, and 45 is used as a multiplier reflecting the average class size at the time.

[29] We thereby also broaden the sample to all men, rather than men who served in WWII—Aaronson and Mazumder (2011) link the census to enlistment records which contain enlistees' county of birth.

TABLE 3: Rosenwald Schools and Education

	Years of Education in 1940 Sample mean: 5.31						
	Baseline	Placebo		Heterogeneity			
		Stayers	Movers	Deep South	Other South	Low HRR	High HRR
Rosenwald exposure	**0.31*****	**0.38*****	**0.09**	**0.66*****	**0.28****	**0.23**	**0.83*****
	(0.10)	(0.12)	(0.23)	(0.19)	(0.13)	(0.15)	(0.22)
Adjusted R^2	0.16	0.16	0.16	0.13	0.17	0.17	0.12
Observations	113,107	90,527	22,520	60,497	52,606	39,125	36,403

Notes: This table shows the effect of Rosenwald school exposure (ranging from 0 to 1) on a person's years of education in 1940. All regressions control for ancestor county FEs, year of birth×ancestor state FEs, and a dummy for the rural/urban status of the part of a county in which a person's ancestors lived in 1870. Columns 2 and 3 compare individuals who remained in the state of their ancestors until at least 1920 ("Stayers") and those who did not ("Leavers"). "Low HRR" considers the three states with the lowest HRR score (least oppressive); "High HRR" those with the highest score (most oppressive). Standard errors are clustered at both the county and cohort level. See Data Appendix A for details on the sample and data.

until at least 1920 benefited from new Rosenwald schools built there after 1919. For families who had moved out of the area by 1920, Rosenwald school in their place of origin had, as expected, no effect on their education (see column 3 of Table 3).

TABLE 4: Intergenerational Effects of the Rosenwald Schools

	Children's Neighborhood Level Outcomes in 2000			
	HS Degree (%) Mean: 69.33	College Degree (%) Mean: 12.15	Income (USD) Mean: 28,831.25	House Value (USD) Mean: 95,986.10
Father's Rosenwald exposure	**2.78**	**4.95****	**2,277.22****	**35,471.16****
	(3.06)	(2.10)	(1,120.35)	(16,917.63)
Level	ZIP×Race×Sex	ZIP×Race×Sex	ZIP×Race	ZIP×Race
Adjusted R^2	0.03	0.03	0.03	0.03
Observations	6,420	6,420	6,434	6,434

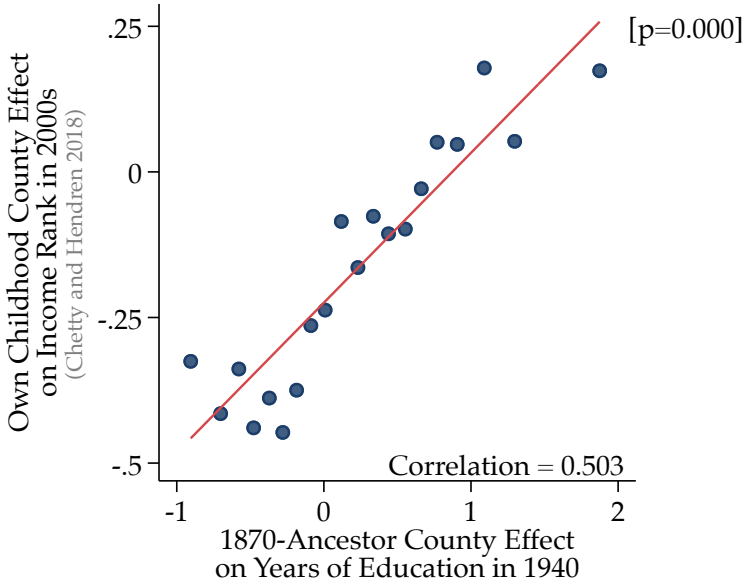
Notes: This table shows the effect of fathers' Rosenwald school exposure (ranging from 0 to 1) on the residential neighborhood quality of the their children in 2000. All regressions control for ancestor county FEs, father's year of birth×ancestor state FEs, and a dummy for the rural/urban status of the part of a county in which a person's ancestors lived in 1870. Standard errors are clustered at both the county and father's cohort level. See Data Appendix A for details on the sample and data.

Attending a Rosenwald school not only increased the education of the benefiting individuals, but continues to affect the economic conditions of their children today (see Table 4). We find that the children of Rosenwald school attendees live in neighborhoods where Black people tend to be more likely to have a college degree, earn higher incomes, and live in more valuable homes. Thus, the Rosenwald school program's positive effects persist beyond the Jim Crow era.

9. The Persisting Geography of Opportunity

In this section, we compare a location's effect on Black economic progress after slavery to the same location's causal effect on intergenerational mobility in recent decades as estimated by Chetty and Hendren (2018). The positive correlation ($\rho = 0.503$) between the historical and modern estimates suggests that a county's ability to generate upward mobility is highly persistent over time. The correlation of effects at the state level is even higher ($\rho = 0.679$, see Appendix Figure J.29). Locations that spurred faster Black economic progress after slavery also increased the intergenerational mobility of American families in recent decades.

FIGURE 16: Persistence of a County's Capacity to Generate Upward Mobility



Notes: This figure shows a binned scatter plot relating the effect of ancestor's county of residence in 1870 on 1940 years of education among descendants of the Enslaved (see Figure 7) to the causal effects those same counties had on intergenerational mobility in recent decades as estimated by Chetty and Hendren (2018). We use their estimates of children's mean percentile rank in the national household income distribution at age 26 conditional on growing up with parents at the 25th percentile of the national household income distribution. We use empirical Bayesian shrinkage as described in Appendix H. Appendix Figure J.29 repeats this figure at the state level. See Data Appendix A for details on the sample and data.

The strength of this correlation may appear surprising given that the institutional and economic environment drastically differed across the two periods. Specifically, throughout 1870 to 1940 the economic progress of Black families in the South was limited by Jim Crow. In addition, the Southern economy was largely agricultural with 40 percent of Black men working in this sector. By the 2000s, Jim Crow had been abolished for more than 30 years and less than two percent of Black men worked in agriculture. However, while Jim Crow ended and the Southern economy evolved, racial oppression did not end. Acharya et al. (2018) show that counties most reliant on slavery until 1865 continue

to have high rates of racial resentment against Black Americans and low political support for policies that could promote Black progress. These factors likely continue to limit Black economic progress in those locations (see also Berger, 2018).

While the effects of places estimated in this paper is specific to the formerly Enslaved, those estimated by Chetty and Hendren (2018) are based on children with parents at the 25th percentile of the income distribution regardless of race. Even if not specifically conditioning on race, the large Black-white disparities in family incomes lead to a disproportionate contribution of Black families to their estimates. We find that the correlation shrinks to less than half ($\rho = 0.204$) when using mobility estimates based on children with parents at the 75th percentile (see Appendix Figure J.30).[30]

10. Conclusion

This paper provides new evidence on the historical roots of modern racial disparities. We find that descendants of the Enslaved face large socioeconomic gaps compared to descendants of the Free even today. This gap accounts for a large fraction of the modern Black-white gaps in education, income, and wealth. Descendants of the Enslaved still face socioeconomic disparities compared to other Black Americans because of their exposure to location-specific factors that slowed economic progress after slavery. Building on the exogenous component in the geographic distribution of the formerly Enslaved, we also find that locations had a large causal effect on their descendants' economic progress until 2000. The most negative effects are concentrated in the Lower South. Those locations were "bad" for race-specific reasons as white Americans did not fare worse there than they did elsewhere. Our evidence is suggestive of education after slavery playing an important role for intergenerational progress. The intensity with which location-specific factors affected Black economic progress intensified as racial segregation increasingly took hold of the Deep South.

This paper has limitations that future work may be able to overcome. First, we limit our analysis to men because automated census-linking methods are either not available or have poor coverage for women. Women have historically tended to change their last names upon marriage, making it impossible for conventional methods to link them across census records. While not within the scope of this paper, in a separate project we link millions of women across census records by using information on their maiden and married names from social security applications (Althoff et al., 2022). Second, while we identify the causal effect of locations on Black economic progress after slavery and show that

[30] We interpret the effect of an ancestor's location of emancipation on their descendant's outcomes in 1940 as a measure of locations' effect on upward mobility. While our estimates do not condition on ancestors' outcomes in 1870 explicitly because of data limitations, they do so implicitly by conditioning on ancestors being enslaved up until 1865.

Jim Crow institutions account for at least half of the variation, we cannot rule out other important factors.

Our findings have important implications for policies that aim to reduce the disadvantage faced by descendants of the Enslaved. First, our results highlight the importance of disparities *within* racial groups that are not necessarily addressed by race-specific policies. College affirmative action is a prime example. Massey et al. (2007) show that the more selective a college, the less likely their Black students descend from the Enslaved. While only 13 percent of the country's 18- to 19-year-old Black Americans have an immigration background, 41 percent of Black Ivy League students do. Affirmative action increases racial diversity on campuses but is less effective in alleviating disadvantages faced by descendants of the Enslaved.

Second, our results highlight that increasing the supply of educational resources to Black children can have long-lasting positive effects on their families' economic progress. We find that externally providing such resources is particularly effective in areas where Black children are most deprived of them. Similar evidence has been documented by development economists (e.g., Duflo, 2001; Glewwe and Kremer, 2006). Our evidence suggests that policies that provide deprived children with educational resources can have large intergenerational spillover effects. Ignoring those effects may lead to the design of policies that are below their optimal scale.

Third, there has been renewed interest in the specific policy of reparations, i.e., wealth transfers to descendants of the Enslaved (e.g., Darity, 2008; Craemer et al., 2020; Boerma and Karabarbounis, 2021). We argue that any assessment of the legacy of slavery should incorporate both when and where a family was freed—i.e., how long they were enslaved and how intensively they were exposed to Jim Crow after slavery. Our empirical evidence suggests that, indeed, Black families continue to be drastically impacted by when and where their ancestors were freed. For that matter, we must stress again that we only quantify the *additional* disadvantage faced by those whose ancestors were enslaved until 1865 and concentrated in the Deep South compared to those who gained freedom earlier, mostly in the Upper South and North. Many in the group of the Free were enslaved in earlier periods, and all Black Americans faced discrimination as a result of slavery regardless of their specific family history. Note that while some argue that reparations should only be received by those who can prove their ancestors were enslaved, our results suggest that post-slavery institutions also harmed Black Americans who descended from the Free—a group that may find it harder to prove their ancestors had been enslaved decades before the Civil War.

REFERENCES

AARONSON, D. AND B. MAZUMDER (2011): "The Impact of Rosenwald Schools on Black Achievement," *Journal of Political Economy*, 119, 821–888.

ABRAMITZKY, R., L. BOUSTAN, E. JACOME, AND S. PEREZ (2021): "Intergenerational Mobility of Immigrants in the United States over Two Centuries," *American Economic Review*, 111, 580–608.

ABRAMITZKY, R., L. BOUSTAN, AND M. RASHID (2020): "Census Linking Project: Version 1.0," dataset: `https://censuslinkingproject.org`.

ABRAMITZKY, R., L. P. BOUSTAN, K. ERIKSSON, J. J. FEIGENBAUM, AND S. PÉREZ (2019): "Automated Linking of Historical Data," Working Paper 25825, National Bureau of Economic Research.

ACEMOGLU, D., S. JOHNSON, AND J. A. ROBINSON (2002): "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*," *Quarterly Journal of Economics*, 117, 1231–1294.

ACHARYA, A., M. BLACKWELL, AND M. SEN (2018): *Deep Roots: How Slavery Still Shapes Southern Politics*.

AGER, P., L. BOUSTAN, AND K. ERIKSSON (2021): "The Intergenerational Effects of a Large Wealth Shock: White Southerners after the Civil War," *American Economic Review*, 111, 3767–3794.

AKBAR, P. A., S. LI, A. SHERTZER, AND R. P. WALSH (2020): "Racial Segregation in Housing Markets and the Erosion of Black Wealth," Tech. rep.

ALEXANDER, M. (2010): *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*.

ALTHOFF, L., H. BROOKES GRAY, AND H. REIHARDT (2022): "Intergenerational Mobility and Assortative Mating," Work in progress.

ARMSTRONG, T. B., M. KOLESŕ, AND PLAGBORG-MØLLER (2021): "Robust Empirical Bayes Confidence Intervals," Tech. rep.

BAKER, R. S. (2022): "The Historical Racial Regime and Racial Inequality in Poverty in the American South," *American Journal of Sociology*, 127.

BAYER, P. AND K. K. CHARLES (2018): "Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940," *Quarterly Journal of Economics*, 133, 1459–1501.

BERGER, T. (2018): "Places of persistence: Slavery and the geography of intergenerational mobility in the United States," *Demography*, 55, 1547–1565.

BERLIN, I. (1974): *Slaves without masters: The free Negro in the antebellum South*, Oxford University Press.

——— (1976): "The Structure of the Free Negro Caste in the Antebellum United States," *Journal of Social History*, 9, 297–318.

BERTOCCHI, G. AND A. DIMICO (2014): "Slavery, education, and inequality," *European Economic Review*, 70, 197–209.

BOERMA, J. AND L. KARABARBOUNIS (2021): "Reparations and Persistent Racial Wealth Gaps," Working Paper 28468, National Bureau of Economic Research.

CARRINGTON, W. J., E. DETRAGIACHE, AND T. VISHWANATH (1996): "Migration with Endogenous Moving Costs," *American Economic Review*, 86, 909–930.

CARSON, S. A. (2008): "The effect of geography and vitamin D on African American stature in the nineteenth century: evidence from prison records," *Journal of Economic History*, 68, 812–831.

CENSUS BUREAU, U. (2017): "Census tract," Available at `https://web.archive.org/web/20170513191843/https://factfinder.census.gov/help/en/census_tract.htm`. Accessed on 14 November 2021.

CHETTY, R. AND N. HENDREN (2018): "The impacts of neighborhoods on intergenerational mobility II: County-level estimates," *Quarterly Journal of Economics*, 133, 1163–1228.

CHETTY, R., N. HENDREN, M. R. JONES, AND S. R. PORTER (2020): "Race and Economic Opportunity in the United States: An Intergenerational Perspective," *Quarterly Journal of Economics*, 135, 711–783.

CHETTY, R., N. HENDREN, AND L. F. KATZ (2016): "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment," *American Economic Review*, 106, 855–902.

CHETTY, R., N. HENDREN, P. KLINE, AND E. SAEZ (2014): "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *Quarterly Journal of Economics*, 129, 1553–1623.

CLARK, G. (2014): *The Son Also Rises: Surnames and the History of Social Mobility*, Princeton University Press.

COLLINS, W. J. AND M. H. WANAMAKER (2022): "African American Intergenerational Economic Mobility since 1880," *American Economic Journal: Applied Economics*.

CONNOR, D. S. AND M. STORPER (2020): "The Changing Geography of Social Mobility in the United States," *Proceedings of the National Academy of Sciences of the United State of America*, 117.

CRAEMER, T., T. SMITH, B. HARRISON, T. LOGAN, W. BELLAMY, AND J. WILLIAM DARITY (2020): "Wealth Implications of Slavery and Racial Discrimination for African American Descendants of the Enslaved," *Review of Black Political Economy*, 47, 218–254.

DARITY, W. (2008): "Forty Acres and a Mule in the 21st Century," *Social Science Quarterly*, 89, 656–664.

DARITY, W., D. HAMILTON, AND K. ZAW (2016): "Race, Wealth and Incarceration: Results from the National Longitudinal Survey of Youth," *Race and Social Problems*, 8, 103–115.

DEBOW, J. D. (1854): *Statistical View of the United States*, A.O.P. Nicholson, Public Printer.

DELL, M. (2010): "THE PERSISTENT EFFECTS OF PERU'S MINING "MITA"," *Econometrica*, 78, 1863–1903.

DELL, M. AND B. A. OLKEN (2019): "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java," *Review of Economic Studies*, 87, 164–203.

DERENONCOURT, E. (2022): "Can You Move to Opportunity? Evidence from the Great Migration," *American Economic Review*, 112, 369–408.

DERENONCOURT, E., C. H. KIM, M. KUHN, AND M. SCHULARICK (2022): "The wealth of two nations: The U.S. racial wealth gap, 1860–2020," Tech. rep.

DONALDSON, D. (2018): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 108, 899–934.

DU BOIS, W. E. B. (1935): *Black Reconstruction in America: An Essay Toward a History of the Part Which Black Folk Played in the Attempt to Reconstruct Democracy in America, 1860–1880*.

DUFLO, E. (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91, 795–813.

FITCHETT, E. H. (1947): "The Status of the Free Negro in Charleston, South Carolina, and His Descendants in Modern Society: Statement of the Problem," *Journal of Negro History*, 32, 430–451.

FOGEL, R. AND S. ENGERMAN (1974): *Time on the Cross. The Economics of American Negro Slavery*.

FONER, E. (1963): *Forever Free: The Story of Emancipation and Reconstruction*, Knopf Doubleday Publishing Group.

FRAZIER, E. F. (1949): *The Negro in the United States.*, The Macmillan Company.

FUJIWARA, T., H. LAUDARES, AND F. VALENCIA C. (2019): "Tordesillas, Slavery and the Origins of Brazilian Inequality," Working paper.

GLEWWE, P. AND M. KREMER (2006): "Schools, Teachers, and Education Outcomes in Developing Countries," Elsevier, vol. 2, chap. 16, 945–1017, 1 ed.

GOLDSTEIN, J. R., M. ALEXANDER, C. BREEN, A. M. GONZÁLEZ, F. MENARES, M. OSBORNE, M. SNYDER, AND U. YILDIRIM (2021): "CenSoc Mortality File: Version 2.0. Berkeley: University of California, 2021." dataset: `https://censoc.berkeley.edu/data/`.

JACKSON, L. P. (1942): *Free Negro labor and property holding in Virginia, 1830-1860*, D. Appleton-Century Company, Incorporated.

JUNG, Y. (2021): "Formation of the legacy of Slavery: Evidence from the US South," Working paper.

KOUSSER, J. M. (1974): *The Shaping of Southern Politics: Suffrage Restriction and the Establishment of the One-Party South, 1880-1910*, Yale University Press.

LOGAN, T. D. (2020): "Do Black Politicians Matter? Evidence from Reconstruction," *Journal of Economic History*, 80, 1–37.

MARGO, R. A. (1991): *Race and Schooling in the South, 1880-1950: An Economic History*, University of Chicago Press.

MASSEY, D. S., M. MOONEY, K. C. TORRES, AND C. Z. CHARLES (2007): "Black Immigrants and Black Natives Attending Selective Colleges and Universities in the United States," *American Journal of Education*, 113, 243–271.

MITCHENER, K. J. AND I. W. MCLEAN (2003): "The Productivity of US States since 1880," *Journal of Economic Growth*, 8, 73–114.

MURRAY, P. (1950): *States' Laws on Race and Color*.

NAIDU, S. (2010): "Recruitment Restrictions and Labor Markets: Evidence from the Post-bellum U.S. South," *Journal of Labor Economics*, 28.

——— (2012): "Suffrage, Schooling, and Sorting in the Post-Bellum U.S. South," Working Paper 18129, National Bureau of Economic Research.

NUNN, N. (2008): *Slavery, Inequality, and Economic Development in the Americas: An Examination of the Engerman-Sokoloff Hypothesis*, Cambridge: Harvard University Press, 148–180.

OLIVETTI, C. AND M. D. PASERMAN (2015): "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940," *American Economic Review*, 105, 2695–2724.

PHILLIPS, U. B. (1918): *American Negro Slavery-A Survey of the Supply, Employment and Control of Negro Labor as Determined by the Plantation Regime*, Louisiana State University Press.

PRITCHETT, J. B. (2001): "Quantitative Estimates of the United States Interregional Slave Trade, 1820-1860," *Journal of Economic History*, 61.

——— (2019): "Demographic Consequences of the Interregional Slave Trade," Working paper.

RANSOM, R. L. AND R. SUTCH (2001): *One Kind of Freedom: The Economic Consequences of Emancipation*, Cambridge University Press.

ROBACK, J. (1984): "Southern Labor Law in the Jim Crow Era: Exploitative or Competitive?" *University of Chicago Law Review*, 51, 1161–1192.

RUGGLES, S., S. FLOOD, R. GOEKEN, J. GROVER, E. MEYER, J. PACAS, AND M. SOBEK (2020): "IPUMS USA: Version 10.0," dataset: `https://doi.org/10.18128/D010.V10.0`.

SACERDOTE, B. (2005): "Slavery and the intergenerational transmission of human capital," *Review of Economics and Statistics*, 87, 217–234.

SCHAMA, S. (2006): *Rough crossings: Britain, the slaves and the American revolution*, Random House.

SMALLWOOD, S. E. (2008): *Saltwater slavery: A middle passage from Africa to American diaspora*, Harvard University Press.

STECKEL, R. H. (1979): "Slave height profiles from coastwise manifests," *Explorations in Economic History*, 16, 363.

STECKEL, R. H. AND N. ZIEBARTH (2016): "Trader selectivity and measured catch-up growth of American slaves," *Journal of Economic History*, 76, 109–138.

TADMAN, M. (1977): "Speculators and slaves in the Old South: a study of the American domestic slave trade, 1820-1860," Ph.D. thesis, University of Hull.

——— (1979): "Slave Trading in the Ante-Bellum South: An Estimate of the Extent of the Inter-Regional Slave Trade," *Journal of American Studies*, 13, 195–220.

——— (2008): "6. The Interregional Slave Trade in the History and Myth-Making of the US South," in *The Chattel Principle*, Yale University Press, 117–142.

TAN, H. R. (2022): "A Different Land of Opportunity: The Geography of Intergenerational Mobility in the Early 20th-Century US," *Journal of Labor Economics*.

USPS (2021): "ZIP Code™ - The Basics," Available at `https://faq.usps.com/s/article/ZIP-Code-The-Basics`. Accessed on 14 November 2021.

WARD, Z. (2020): "The Not-So-Hot Melting Pot: The Persistence of Outcomes for Descendants of the Age of Mass Migration," *American Economic Journal: Applied Economics*, 12, 73–102.

WESTERN, B. (2006): *Punishment and Inequality in America*, Russell Sage Foundation.

WOLF, E. S. (2006): *Race and Liberty in the New Nation: Emancipation in Virginia from the Revolution to Nat Turner's Rebellion*, LSU Press.

WOODSON, C. G. (1918): *A Century of Negro Migration*.

WOODWARD, C. V. (1955): *The Strange Career of Jim Crow*, Oxford University Press.

Appendix

A Data Appendix **54**

A.1 Census Data 54

A.2 Census Linking 54

A.3 Identifying Descendants of the Free and the Enslaved 54

A.4 Main Sample 62

A.5 Individual-Level Outcome Variables 62

A.6 Neighborhood-Level Outcome Variables 64

A.7 County Characteristics 66

A.8 Nine-Digit ZIP to Census 2000 Crosswalks 67

B Interpreting the Free-Enslaved Gap **72**

B.1 Selection into Being Free 73

B.2 Bounding the Selection Bias 74

C Socioeconomic Status among Descendants of the Free and the Enslaved **81**

D Migration among Descendants of the Free and the Enslaved **82**

E Adjusting Estimates for Misclassification Bias **86**

F Adjusting Estimates for Intermarriage **87**

G The Direct Effect of Locations After Accounting for Migration **89**

H Empirical Bayes Shrinkage of Place Effects **90**

I Placebo Exercises **93**

J Additional Results **99**

J.1 The Free-Enslaved Gap in the 21st Century 99

J.2 The Free-Enslaved Gap in Alternative Measures 100

J.3 The Free-Enslaved Gap in Intergenerational Mobility 100
J.4 The Free-Enslaved Gap Based on the Distribution of Last Names 102
J.5 The Free-Enslaved Gap for Free Without Physical and Human Capital . . 102
J.6 The Free-Enslaved Gap between and within Ancestor's Birthplace 103
J.7 The Free-Enslaved Gap in Mortality 106
J.8 Decomposing the Black-White Gap in Education 106
J.9 The Intergenerational Effect of Exposure to State-Specific Factors 107
J.10 The Persistence of Place Effects 111
J.11 Additional Results from the Border Discontinuity Design 112
J.12 Cohort-Specific Effects of Exposure to Regional Factors 116

A. Data Appendix

A.1 Census Data

We use the publicly available full-count U.S. Census Bureau data at the individual level for Black Americans for all decades between 1850 and 1940. The 1850 census is the first to include individual-level data; the 1940 census is the latest for which restricted access to the names of individuals is available. We use linking methods that rely on those names. The data from 1890 were destroyed by a fire and are therefore not included.

A.2 Census Linking

We use crosswalks between observations in different census decades provided by Abramitzky et al. (2020), publicly available here at `https://censuslinkingproject.org`. The crosswalks can be merged into the public version of the census data using the histid identifier. We do so, linking all adjacent and non-adjacent census decades in our sample period.

There are multiple crosswalks available, each based on different linking techniques. Our main results use the "abe_race_nysiis_standard" link, which matches observations based on first name, last name, and age. It requires each name to be unique within a five-year window for each race but allows some names to be matched even if their spelling differs. We also provide results for "abe_nysiis_conservative," which requires names to be unique within and across races and allows for fewer deviations in matching characteristics.

To study intergenerational dynamics, we inherit parents' characteristics to the children in their household. Certain characteristics, such as occupation or education, are only inherited from prime-age male (ages 20–54) ancestors to ensure comparability over time.

A.3 Identifying Descendants of the Free and the Enslaved

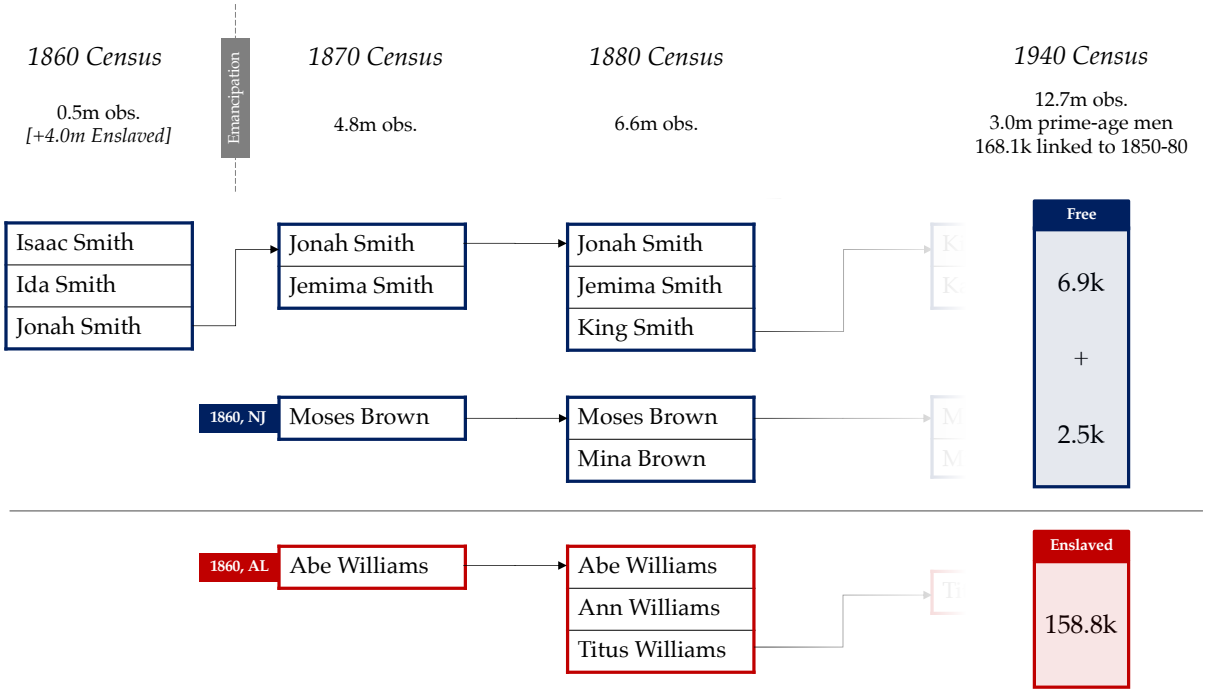
Main Method: Linking Historical Census Records

Figure A.1 illustrates our new method to identify descendants of the Free and descendants of the Enslaved in census records between 1870 and 1940. It mainly relies on census-linking methods (Abramitzky et al., 2019) but also uses information on place and year of birth.

The method consists of three steps. First, we identify the Free themselves before identifying their descendants. In 1850 and 1860, the enslaved population was excluded from the individual-level censuses. By definition, every Black American included in the census was therefore free before 1865. We link the 1850 and 1860 censuses forward to all

census decades between 1870 and 1940 and then classify every Black American who can be linked to 1850 or 1860 as free.

FIGURE A.1: Illustration of Our Free-Enslaved Classification Algorithm



Notes: This figure illustrates our new method to identify descendants of the Free and the Enslaved in census records 1870-1940. The names are chosen are arbitrary examples and do not reflect real data. Jonah Smith is identified as a descendant of the Free because he can be linked back to the 1860 census; Moses Brown because he was born in a state (New Jersey) that had abolished slavery by the time of his birth (1860). Abe Williams does not fall into either category and is therefore classified as formerly enslaved or a descendant of the Enslaved. The Free-Enslaved status is assigned to descendants based on their male ancestor. In 1940, the final year of our sample, we identify 9,400 descendants of the Free (6,800 through direct linking to 1850–1860 and 2,600 through their ancestor's birthplace) and 155,800 descendants of the Enslaved. While not comprehensively illustrated here, we do link across all adjacent and non-adjacent census records of 1850, 1860, 1870, 1880, 1900, 1910, 1920, 1930, and 1940.

In addition to linking, we use information on place and year of birth in our classification algorithm. All Northern states had begun banning or restricting slavery by 1804—some of them decades earlier. Any Black person born in those states was either free upon birth or would be emancipated by a certain age (typically in their 20s). While the latter case opens up the possibility of a Northern-born Black person being sold into slavery in other states before their emancipation, this possibility was ruled out by law.

In Appendix Table A.2, we compare the de jure to the de facto status of slavery in the North. As a de facto measure, we show the number of slaves in the state in absolute numbers and as a fraction of the state's Black population. Based on this evidence, we classify any Black American born outside of the slave states after 1804 and before 1865 as Free. In addition, we use the state-specific years in which slavery was abolished or restricted in non-slave states to go even further back in time.

Second, we identify the *descendants* of the Free by using information on the relationship between individuals within census households. Specifically, we classify Black people with a free Black American *ancestor* as being descendants of the Free. Any person without a free ancestor is classified as a descendant of the Enslaved. In 1940, the final year of our sample, we identify 9,400 descendants of the Free and 155,800 descendants of the Enslaved.

Attenuation bias. Because we can only link men, the descendant classification is determined exclusively through male ancestors. This data limitation prevents us from assessing inter-marriage between the Free and the formerly Enslaved or their descendants. The fact that some Black Americans will be both descendants of the Free and descendants of the Enslaved potentially biases our estimates of the Free-Enslaved gap toward zero. This attenuation bias is one reason why our results should be interpreted as a lower bound for the true Free-Enslaved gap.

Linking bias. Any study that uses automated linking methods faces the problem that individuals who can be linked across decades may not represent the overall population. For example, families with a high socioeconomic status may choose more unique names for their children, making it easier to create a unique match across census records. A socioeconomic gap between two sub-populations is only biased if the linking procedure differentially selects them into the sample. Table A.1 shows that, if anything, the linking procedure biases the Free-Enslaved gap toward zero.

TABLE A.1: Assessing Linking Bias

	Free (1860)			Enslaved (1870)		
	Linked	Population	Δ	Linked	Population	Δ
Literacy (%)	65.1	66.8	-3%	20.4	20.4	0%
Occupation Score	6.0	6.1	-1%	3.7	3.8	-1%
Real property ($)	1,217	1,230	-1%	1,400	1,270	10%
Personal property ($)	312	316	-1%	312	293	6%
Lives in North (%)	45.1	52.1	-13%	7.8	8.2	-4%
Lives on Farm (%)	21.2	18.2	17%	23.8	23.2	3%
Observations	20,994	79,374		190,676	726,667	

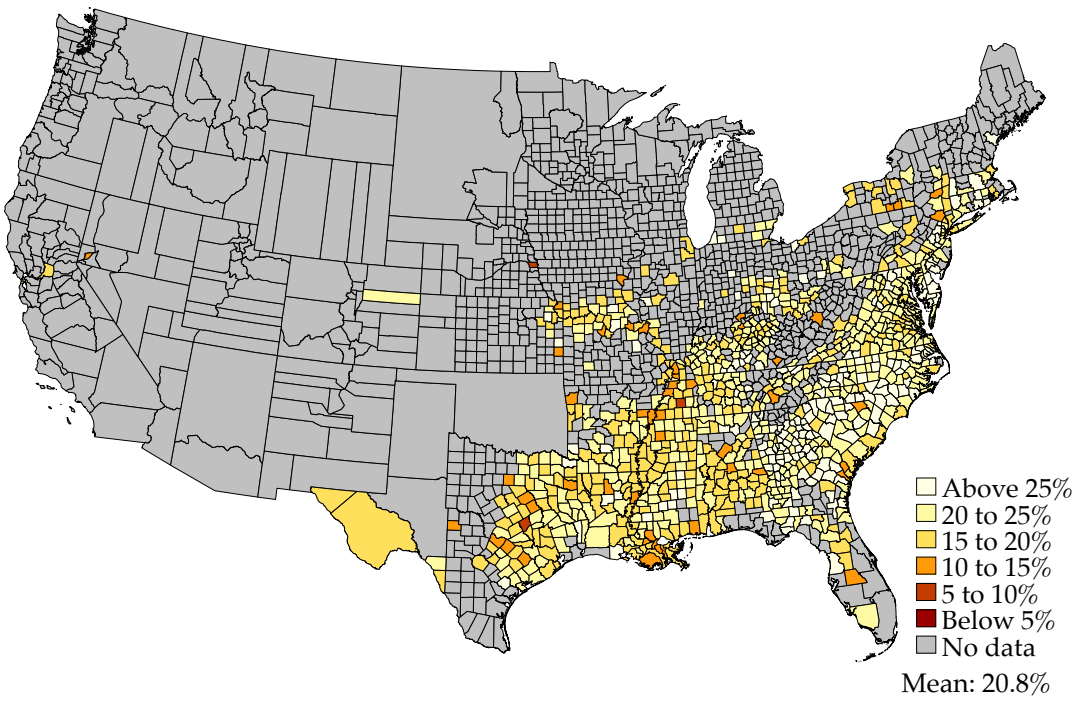
Notes: This table suggests that sample selection is small and negative among the Free and more positive among the formerly Enslaved. If anything, this kind of selection biases the Free-Enslaved gap toward zero. The left panel compares the Free who can be linked to any future decade to the entire population (which only consists of free Black Americans) in 1860. The right panel compares the formerly Enslaved (anyone not classified as Free by our algorithm) in 1870 who can be linked to any future decade to the 1860 population of the formerly Enslaved.

In addition, a family's socioeconomic status may affect not only *whether* they can be linked across decades but also *over how many decades* they can be linked. For example, children who grow up with single mothers can typically not be linked to their grandparents because women cannot be linked by these methods due to name changes at mar-

riage. Our classification algorithm identifies descendants of the Free mainly through whether they can be linked back to 1850 or 1860, which could lead to an almost mechanically higher socioeconomic status. We addressed this concern in Section 3.3.2 (see Figure 1).

Last, one may be concerned that the effect of place in 1870 on outcomes in 1940 may be biased by differences in linking rates across those locations. In particular, areas with large Black populations may have lower linking rates because the linking relies on the uniqueness of a person's name, state of birth, and age. Lower linking rates may imply that only individuals with particularly rare names—and therefore potentially different socioeconomic statuses—are selected into the sample. Figure A.2 addresses this concern by showing counties' average likelihood of a resident in 1870 being linkable to any future decade. Unsurprisingly, linking rates are somewhat higher in the Upper South and North, where the Black population tended to be smaller in 1870. If anything, this would imply that the positive selection into linking illustrated in Table A.1 would further bias the Free-Enslaved gap toward zero.

FIGURE A.2: Linking Rates by County



Notes: This figure shows the average linking rate for Black prime-age (20–54) men in the 1870 to any future decade (1880, 1900, 1910, 1920, 1930, or 1940). Only counties with a Black population of at least 100 prime-age men in 1870 are included.

TABLE A.2: Abolition of Slavery in the North

Year	State	*De Jure* Abolition of Slavery	*De Facto* Number of Slaves	
			Year	Total
1777	Vermont	Slavery was banned immediately upon founding of Vermont (Constitution of Vermont, 1777).	1790	0[31]
			1800	0
			1810	0
			1820	0
			1830	0
			1840	0
			1850	0
1780	Pennsylvania	Law of gradual emancipation passed in 1780 (Pennsylvania General Assembly, 1780). Black Americans born to enslaved mothers after 1780 would be freed at age 28. Slavery was ended in 1847.	1790	3,737 (36%)
			1800	1,706 (10%)
			1810	795 (3%)
			1820	211 (1%)
			1830	403 (1%)
			1840	64 (0%)
			1850	0
1781	Maine Massachusetts	Slavery was abolished by Supreme Judicial Court rulings in three related court cases, collectively known as the "Quock Walker case" (Cushing, 1961; Zilversmit, 1968). Slavery was ruled incompatible with the new state constitution of 1780.	1790	0
			1800	0
			1810	0
			1820	0
			1830	3 (0%)
			1840	0
			1850	0
1783	New Hampshire	Similar to Massachusetts, New Hampshire's constitution essentially abolished slavery by stating "all men are born equal and independent" (Constitution of the State of New Hampshire, 1783). However, it is not clear whether court rulings indeed interpreted the constitution as being at odds with slavery or not.	1790	158 (20%)
			1800	8 (1%)
			1810	0
			1820	0
			1830	3 (0%)
			1840	1 (0%)
			1850	0
1784	Rhode Island	Law for gradual emancipation passed in 1784 (General Assembly of Rhode Island, 1784). Black Americans born to enslaved mothers after 1784 would be freed at age 18 (women) or 21 (men).	1790	952 (22%)
			1800	381 (10%)
			1810	108 (3%)
			1820	48 (1%)
			1830	17 (0%)
			1840	5 (0%)
			1850	0

TABLE A.2: Abolition of Slavery in the North

Year	State	*De Jure* Abolition of Slavery	*De Facto* Number of Slaves	
			Year	Total
1784	Connecticut	Law for gradual emancipation passed in 1784 (Connecticut General Assembly, 1784). Black Americans born to enslaved mothers after 1784 would be freed at age 25. This age was lowered to 21 in 1797. Slavery was abolished in 1848.	1790	2,759 (50%)
			1800	951 (15%)
			1810	310 (5%)
			1820	97 (1%)
			1830	25 (0%)
			1840	17 (0%)
			1850	0
1787	Ohio Indiana Illinois Michigan Wisconsin Minnesota	The Confederation Congress's Northwest Ordinance of 1787 both banned and enforced slavery (Confederation Congress, 1787). A clause allowed Northerners to capture and enslave runaway slaves. Slavery was abolished by Ohio in 1802, Indiana in 1816, and Illinois in 1818.	1790	–
			1800	135 (21%)
			1810	429 (28%)
			1820	1,106 (40%)
			1830	788 (5%)
			1840	348 (1%)
			1850	0
1799	New York	Law for gradual emancipation passed in 1799 (New York State Legislature, 1799). Black Americans born to enslaved mothers after 1799 would be freed at age 25 (women) or 28 (men). In 1817, state decided to free all slaves born before 1799 (but not their children) in 1827 (New York State Legislature, 1817).	1790	21,324 (82%)
			1800	20,343 (66%)
			1810	15,017 (37%)
			1820	10,088 (26%)
			1830	75 (0%)
			1840	4 (0%)
			1850	0
1804	New Jersey	Law for gradual emancipation passed in 1804 (New Jersey State Legislature, 1804). While not freeing living slaves, Black Americans born to enslaved mothers after 1804 would be freed at age 21 (women) or 25 (men).[32]	1790	11,423 (81%)
			1800	12,422 (74%)
			1810	10,851 (58%)
			1820	7,557 (38%)
			1830	2,254 (11%)
			1840	674 (3%)
			1850	236 (1%)

Notes: This table provides a timeline for the abolition of slavery in the North. The first column indicates the year which we choose as the states' final year of slavery. We classify any Black American born in the state after this cutoff as free. The third column shows the laws that abolished slavery. In many cases, slavery was not abolished outright, but rather it was restricted in ways that would imply a person is free before 1865 in all likelihood. The final column shows the actual number of slaves who reside in the state and the percentage of the state's Black population being enslaved in parentheses. The number of slaves is taken from aggregate counts in census records (1790–1850).

[32]While the 1790 census states that 16 slaves were in Vermont that year, this is likely an error.

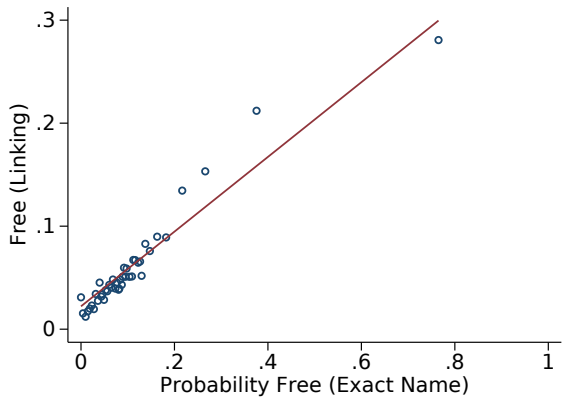
[32]There is some evidence that after 1804, some Black Americans were sold to slave states before they

Alternative Method of Free-Enslaved Classification: Distribution of Last Names

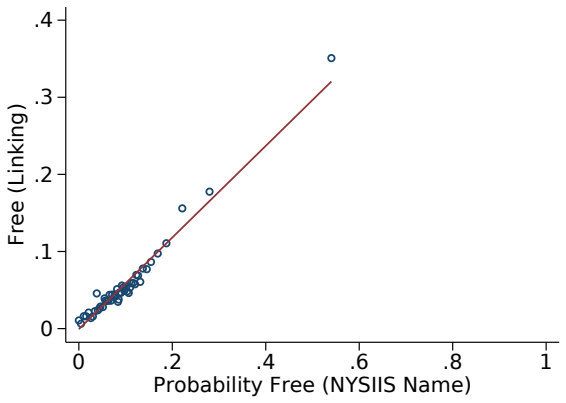
While our main method provides a high-accuracy classification of descendants of the Free and the Enslaved, accuracy comes at the cost of reduced sample sizes due to imperfect linking rates across the decades. To use the full census sample of Black Americans after 1870, rather than a linked sub-sample thereof, we develop an additional strategy for identifying descendants of the Free and the Enslaved based on last names. Figure A.3 shows that the name-based measures are highly correlated with the free status based on our preferred measure, though they are attenuated as expected.

FIGURE A.3: Comparing Name-Based and Linking-Based Measures

(A) Exact Last Names



(B) Phonetic Version of Last Names



Notes: This figure compares the probabilistic measures of descending from free Black Americans with our preferred measure based mainly on census linking. This binned scatter plot shows that among Black prime-age men in the 1940 census, the fraction of people classified as Free closely coincides with the predicted probability based on the people's last names.

Our alternative classification algorithm uses changes in the distribution over last names from 1850–1860 to 1870–1880. Before 1865, the census only includes free Black Americans—after, it also includes the formerly Enslaved and their descendants.

We compute the relative frequency of each last name before and after 1865. We then create a measure of how likely a person is to descend from the Free by dividing their last name's relative frequency before 1865 by its relative frequency after 1865. For example, the last name Du Bois appears with relatively high frequency in the 1850 and 1860 censuses, while Freedman does not appear at all. After the four million formerly enslaved individuals enter the census sample in 1870 and 1880, the name Du Bois is far less (one-tenth) frequent, whereas a substantial number of individuals entered the sample with the last name Freedman for the first time. These changes suggest that anyone named Du Bois after 1865 likely descends from the Free, whereas anyone named Freedman likely descends from the Enslaved. Note that not all names give us a good idea of whether a

reached the age to be emancipated (Armstead et al., 2016, p.104).

person descends from the Enslaved or not. Especially names very common among Black Americans before 1865, such as Johnson, Brown, or Smith, remain very common after 1865. Other names such as Washington did exist among Black Americans before 1865, but became substantially more common after many newly freed enslaved people chose this name in honor of the country's first president.

Formally, using the example of the last name Du Bois, we estimate the name-specific likelihood of descending from free Black Americans defined as

$$\begin{aligned} P(Free_{it} = 1 | Name_i = \#DuBois_t) &= \frac{P(Free_{it} = 1, Name_{it} = \#DuBois_t)}{P(Name_{it} = \#DuBois_t)} \\ &= \frac{P(Free_{i,1860} = 1, Name_{i,1860} = \#DuBois_t)}{P(Name_{i,1870} = \#DuBois_t)} \\ &= \frac{P(Name_{i,1860} = \#DuBois_t)}{P(Name_{i,1870} = \#DuBois_t),} \end{aligned}$$

where the second equation follows from assuming that a last name conveys a constant probability of descending from free Black Americans. The last equation follows from the fact that the 1860 census only contained free Black Americans. This equation can be approximated by

$$\widehat{P}(Free_{it} = 1 | Name_{it} = \#DuBois_t) = \frac{\#(\#DuBois_t)_{1860} / BlackPop_{1860}}{\#(\#DuBois_t)_{1870} / BlackPop_{1870}},$$

where $\#DuBois_t$ is the number of individuals with the last name Du Bois in a given year and $BlackPop_t$ is the population of all Black Americans (free and enslaved). To reduce noise, we combine the names from the 1850 and 1860 censuses as a pre-1865 count and the 1870 and 1880 censuses as a post-1865 count. Before 1865, we compute the population by adding up the census sample size (the Free) and the number of the Enslaved (Berlin, 1974). We truncate our estimated probability by 0 and 1. Names that only appear pre-1865 but not post-1865 are assigned probability 1; those that only appear post-1865 are assigned probability 0.

To allow for misspellings, we also compute this measure based on the phonetics of last names. Specifically, we transform last names using the New York State Identification and Intelligence System (NYSIIS) phonetic code. For example, the last names "Browne" and "Brown" both become "Bran." For placebo exercises, we also compute the above measure as a pseudo-probability of being free for white Americans as well as for 1875 as a time placebo for Emancipation.

A.4 Main Sample

For our main sample, we focus on prime-age (20–54) Black men who can be linked to ancestors in 1880 or before. Our focus on prime-age individuals provides a certain form of comparability, limiting the possibility that an individual is in school or retired. We focus on men because we rely on automated census-linking techniques that are either unavailable or have notoriously low coverage for women. We restrict the sample to individuals who can be linked back to 1880 or before for two reasons. First, this requirement excludes families who migrated to the U.S. after 1880. Any comparison made in our analysis will thus be for individuals whose families have lived in the country for at least 60 years. Second, it reduces the potential for linking bias as discussed in Section A.3.

A.5 Individual-Level Outcome Variables

Our main outcomes variables can be categorized as (proxies of) income, education, or wealth. Most individual-level data draw on census records provided through IPUMS (Ruggles et al., 2020).

Income

- **Occupational income scores, 1850–1940 (census).** Because the census does not include any continuous measure of income before 1940, researchers have instead relied on occupational income scores. The most popular version, "occscore," reflects the median total income of a person in that occupation in 1950.

- **Lido income scores, 1850–1940 (Saavedra and Twinam, 2020).** Occupational income scores do not contain any age-, sex-, or race-specific information. The recent literature has used regression and machine learning techniques to improve on the traditional occupational income score (e.g., Saavedra and Twinam, 2020; Abramitzky et al., 2021). We use the Lido score constructed by Saavedra and Twinam (2020). The authors constructed it using machine learning techniques using 1950 and 2000 census data to validate their results against occscore in the 1915 Iowa census. According to Abramitzky et al. (2021), the Lido score has a correlation of 0.99 with their own measure.

- **Occupational skill, 1850–1940 (Leeuwen and Maas, 2011).** We use HISCLASS, a classification to compare occupations based on the skill they typically required. The classification ranges from "higher managers" to "unskilled farm workers." We coarsen this classification by assigning "skilled" to every occupation classified as "medium skilled workers" or above and "unskilled" to everyone else.

- **Wage income, 1940 (census).** We use wage income for 1940, the only year it is available for in our sample period.

- **Predicted total income, 1940 (census).** Total income, the sum of wage and non-wage income, only becomes available in the census after 1940. Following the prior literature, we use a rich set of covariates to predict total income in 1940 (see, e.g., Abramitzky et al., 2021). In particular, we make use of the fact that in 1960, the census includes both wage income and total income. We regress 1960 total income on 1960 wage income, a quadratic function in age, and all interactions of state fixed effects $\times$ self-employment status $\times$ occupation scores. To avoid over-fitting, we shrink the coefficients using LASSO with tenfold cross-validation. We then apply the estimated equation to 1940, yielding a prediction of 1940 total income.

  Note that we use 1960 data for the prediction because 1950 is affected by the "sample-line problem," reducing the number of observations that include information on incomes.

Education

- **Literacy, 1850–1940 (census).** We use literacy for all years. In 1940, literacy becomes unavailable, and instead the census starts to include educational attainment. We proxy for literacy by having completed at least the second grade.

- **Years of education, 1940 (census).** We impute years of education from the highest educational level attained ("educd").

- **High school, 1940 (census).** We impute whether a person holds a high school degree based on whether they completed at least 12 years of schooling ("educd").

- **College, 1940 (census).** We impute whether a person holds a college degree based on whether they completed at least 16 years of schooling ("educd").

- **Graduate, 1940 (census).** We impute whether a person holds a graduate degree based on whether they completed at least 17 years of schooling ("educd").

Wealth

- **Personal property, 1860–1870 (census).** Measures "the contemporary dollar value of all stocks, bonds, mortgages, notes, livestock, plate, jewels, and furniture" as reported to the census. It is not clear whether zeros indicate missing values or true zero personal property, and therefore we replace zeros with "missing."

- **Real property, 1850–1870 (census).** Measures "the contemporary dollar value of any real estate owned by the respondent" as reported to the census. It is not clear

whether zeros indicate missing values or true zero personal property, and therefore we replace zeros with "missing."

- **Homeownership, 1850–1940 (census).** Measures whether the individual rents or owns their home. For 1900 to 1940, the census reports homeownership directly. For 1850 to 1870, we follow Collins and Margo (2011) in imputing homeownership status using information on wealth, where every household with positive real property is classified as owner-occupied. Collins and Margo (2011) exempt households who live in multi-family homes from this classification but the information necessary to follow them in doing so is not included in the full-count version of the census we use. However, creating homeownership proxies using their and our method yields a correlation of 0.9733 in the 1 percent sample.

- **House value, 1930–1940 (census).** Measures the house value conditional on owning the house.

A.6 Neighborhood-Level Outcome Variables

While we cannot link our data to censuses after 1940, we can link the 1940 census to administrative death records from 1988 and 2005 using the CenSoc-Numident file (Goldstein et al., 2021). Importantly, the death records contain the nine-digit ZIP codes of residence at the time of death. We link these codes to statistical census geographic areas, i.e., census tracts, block groups, and blocks (see Section A.8 for more detail on the procedure). Census tracts contain between 1,200 and 8,000 people and are designed to be "relatively homogeneous units with respect to population characteristics, economic status, and living conditions" (Census Bureau, 2017). Block groups (between 600 and 3,000 people) and blocks are subdivisions of a census tract.

We assigned to each decedent various socioeconomic characteristics based on these statistical areas at the time of death. Since the sample is about evenly split between deaths before 2000 and deaths after 2000, we used the aggregated census data for the year 2000 from the NHGIS database. For variables from other sources, we selected the data to refer to a period as close to 2000 as availability allowed.

Income

- **Income, 2000 (NHGIS).** The median household income by race of householder. Available by ZCTA, census tracts, and block groups.

Wealth

- **House value, 2000 (NHGIS).** The median value of owner-occupied housing units by race of householder. Available by ZCTA and census tracts.
- **Homeownership, 2000 (NHGIS).** The share of occupied housing units that is occupied by the owner (relative to a renter) by race. Available by ZCTA, census tracts, block groups, and blocks.

Education

- **High school degree, 2000 (NHGIS).** The share of the population over 25 years old by race and sex who hold a high school degree. Available by ZCTA, census tracts, and block groups.
- **College degree, 2000 (NHGIS).** The share of the population over 25 years old by race and sex who hold a college degree. Available by ZCTA, census tracts, and block groups.

Health

- **Age at death, 1988–2005 (BUNMD, Goldstein et al., 2021).** The median age at death by race and sex. Available by five-digit ZIP code, census tracts, block groups, and block.
- **Life expectancy, 2010–2015 (National Center for Health Statistics, 2021).** Estimates of life expectancy at birth. Available by census tracts.
- **Physical health, 2017–2018 (CDC, 2020).** Prevalence of poor physical health in the last 14 days among individuals over 18 years. Available by ZCTA and census tracts.
- **Mental health, 2017–2018 (CDC, 2020).** Prevalence of poor mental health in the last 14 days among individuals over 18 years. Available by ZCTA and census tracts.

Miscellaneous

- **Evictions, 2000–2005 (Desmond et al., 2018).** Number of eviction filings and evictions per 100 renter-occupied households. Available by census tracts and block groups.
- **Fatal police encounters, 2000–2021** (`fatalencounters.org`). "Fatal Encounters documents non-police deaths that occur when police are present or are precipitated by police action or presence. Officer deaths are included when caused by another

officer, including friendly fire incidents, and criminal actions—like domestic violence—and suicides that occur when other officers are present. Officer vehicle-related deaths are included when they are caused by another officer. Homicides of officers by felons or deaths in the regular course of duties are not generally documented in the database." The data cover the entire U.S. from 2000 to September 09, 2021, but the database is continuously updated by journalist D. Brian Burghart. Available by five-digit ZIP code.

- **Racial segregation, 2000 (NHGIS)**. The Theil Index of racial segregation using the racial composition of a census tract or block group relative to the block that it contains (Theil, 1972). For the exact formula, see Equation (4) in Chetty et al. (2014). Whereas Chetty et al. (2014) compute the measure on the level of the commuting zone relative to the census tracts it contains, we compute the measure on the level of a census tract and block group relative to the block. Available by census tracts and block groups.

- **Percentage Black, 2000 (NHGIS)**. The share of the population that is Black. Available by ZCTA, census tracts, block groups, and blocks.

A.7 County Characteristics

To assess the characteristics of "good" and "bad" counties and the persistence over time, we compile a dataset on the county level. (Manson et al., 2021).

- **Distance to the North, East (NHGIS).** A county's distance to the North and the East is proxied by its centroid's latitude and longitude.

- **Free, 1860 (NHGIS).** Measures the percentage of a county's 1860 Black population that is free.

- **Black, 1860 (NHGIS).** Measures the percentage of a county's 1860 population that is Black.

- **Tobacco, cotton, rice, and sugar, 1860 (NHGIS).** Measures the value of a county's tobacco, cotton, rice, or sugar output as a percentage of the total agricultural output in 1860.

- **Population density, 1870 (NHGIS).** Measures a county's 1870 population per square kilometer area.

- **School, 1870 (NHGIS).** Measures the fraction of a county's Black children (ages 6–16) attending school in 1870.

- **Farm, 1870 (NHGIS).** Measures the fraction of a county's population living on a farm in 1870.

- **Migration cost North, 1870 (Donaldson and Hornbeck, 2016).** Measures the transportation cost through land and water ways from a given county to the Northern cities that were the main destinations of the Great Migration: Chicago, Detroit, Pittsburgh, and New York. The migration cost estimates are based on the 1870 railroad network.

- **Lynchings, 1883–1941 (Seguin and Rigby, 2019).** Measures the number of lynchings that occurred in a county between 1883 and 1941.

- **Racial segregation, 1880 and 1940 (Logan and Parman, 2017).** Measures racial segregation based on a comparison of the probability of different-race neighbors in a county relative to the counterfactual probability had the population been randomly distributed across the county.

- **Slaves per slaveholder, 1860 (NHGIS)**. The average number of enslaved people per slaveholder.

- **Intergenerational mobility, 1996–2012 (Chetty and Hendren, 2018).** Measures the causal effect of a county on the expected rank in the national income distribution conditional on one's parents' income ranking at the 25th percentile during childhood.

A.8 Nine-Digit ZIP to Census 2000 Crosswalks

The administrative death records contain nine-digit ZIP codes ("ZIP9") of the place of residence at the time of death. We use the Census Bureau's TIGER/Line ASCII files (1994, 1995, 1997, 1998, 1999, 2000, 2002, 2003, 2005, and 2006) to link ZIP9s to 2000 census statistical areas (i.e., census blocks, block groups, and census tracts). A ZIP9 is a characteristic of a range of addresses, usually a side or segment of a street. The relevant records in the TIGER/Line files for our purpose are record types 1, 6, and Z.[33] Each entry in record type 1 represents a complete chain (a street segment) and contains the five-digit ZIP ("ZIP5") for the main address range of the complete chain. It also contains the census block number of the polygon on either side of the complete chain.

Record type 6 provides remaining address ranges and their ZIP5s in case the relevant segment of a street is associated with multiple address ranges. Record type Z provides ZIP+4 add-on codes for each address range in record types 1 and 6. Merging the three record types, we obtain a database of ZIP9s and corresponding blocks. The TIGER/Line

[33] Extensive documentation is available online for each TIGER/Line version.

versions before 2000 linked ZIP9s to 1990 census areas. For those versions, we extract the ZIP9 and longitude and latitude of the beginning and end of the street segment that the complete chain corresponds with. Then, we map the street segment's midpoint onto the 2000 census shape files.

In most cases, a ZIP9 maps into a unique block (and hence maps into a unique block group and census tract). For instance, in 2000, 81 percent of ZIP9s were matched to a unique block. For block groups and census tracts, 96 percent and 97 percent of the ZIP9 matches were unique, respectively. In cases where a ZIP9 occurs in more than one statistical area, we assign the area that has the largest number of matches with the relevant ZIP9. This yields a one-to-one mapping of ZIP9s to blocks for each TIGER/Line year between 2000 and 2006. However, not all ZIP9s in the Censoc-Numident death records occur in the TIGER/Line files. To improve the coverage, we sort the data by ZIP9 for each version and interpolate the census statistical areas in case the next non-missing census area is exactly equal to the previous non-missing area (using that the ZIP9s are ordered geographically).

Last, for each decedent, we assign the census area corresponding to their ZIP9 derived from a TIGER/Line version before and after their year of death (if available). For instance, if someone was born in 1996, we first try to assign the census area based on the TIGER/Line in 1995 and in 1997. If either of them is not available, we try to match using the next proximate version. Using this procedure, we link around 84 percent of the decedents with ZIP9s to a census tract, 82 percent to a block group, and 77 percent to a block. For decedents for which we can find the census area corresponding to their ZIP9 both before and after their death, the agreement rate between the different versions is high (98 percent for census tracts, 96 percent for block groups, and 88 percent for blocks).

REFERENCES

ABRAMITZKY, R., L. BOUSTAN, E. JACOME, AND S. PEREZ (2021): "Intergenerational Mobility of Immigrants in the United States over Two Centuries," *American Economic Review*, 111, 580–608.

ABRAMITZKY, R., L. BOUSTAN, AND M. RASHID (2020): "Census Linking Project: Version 1.0," dataset: https://censuslinkingproject.org.

ABRAMITZKY, R., L. P. BOUSTAN, K. ERIKSSON, J. J. FEIGENBAUM, AND S. PÉREZ (2019): "Automated Linking of Historical Data," Working Paper 25825, National Bureau of Economic Research.

ARMSTEAD, S., B. SUTTER, P. WALKER, AND C. WIESNER (2016): ""And I Poor Slave Yet": The Precarity of Black Life in New Brunswick, 1766–1835," in *Scarlet and Black:*

Slavery and Dispossession in Rutgers History, ed. by M. J. Fuentes and D. G. White, Rutgers University Press, 91–122.

BERLIN, I. (1974): *Slaves without masters: The free Negro in the antebellum South*, Oxford University Press.

CDC (2020): "PLACES: Local Data for Better Health, Census Tract Data 2020 release," dataset: https://chronicdata.cdc.gov/500-Cities-Places/PLACES-Local-Data-for-Better-Health-Census-Tract-D/cwsq-ngmh.

CENSUS BUREAU, U. (2017): "Census tract," Available at https://web.archive.org/web/20170513191843/https://factfinder.census.gov/help/en/census_tract.htm. Accessed on 14 November 2021.

CHETTY, R. AND N. HENDREN (2018): "The impacts of neighborhoods on intergenerational mobility II: County-level estimates," *Quarterly Journal of Economics*, 133, 1163–1228.

CHETTY, R., N. HENDREN, P. KLINE, AND E. SAEZ (2014): "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *Quarterly Journal of Economics*, 129, 1553–1623.

COLLINS, W. J. AND R. A. MARGO (2011): "Race and Home Ownership from the Civil War to the Present," *American Economic Review: Papers and Proceedings*, 101, 355–359.

CONFEDERATION CONGRESS (1787): "An Ordinance for the Government of the Territory of the United States North West of the River Ohio," Library of Congress: https://lccn.loc.gov/90898154.

CONNECTICUT GENERAL ASSEMBLY (1784): "An Act Concerning indian, Molatto and Negro Servants and Slaves," Western CT State University: https://archives.library.wcsu.edu/omeka/items/show/2625.

CONSTITUTION OF THE STATE OF NEW HAMPSHIRE (1783): https://www.nh.gov/glance/constitution.htm.

CONSTITUTION OF VERMONT (1777): https://avalon.law.yale.edu/18th_century/vt01.asp.

CUSHING, J. D. (1961): "The Cushing Court and the Abolition of Slavery in Massachusetts: More Notes on the "Quock Walker Case"," *American Journal of Legal History*, 5, 118–144.

DESMOND, M., A. GROMIS, L. EDMONDS, J. HENDRICKSON, K. KRYWOKULSKI, L. LEUNG, AND A. PORTON (2018): "Eviction lab national database: Version 1.0," .

DONALDSON, D. AND R. HORNBECK (2016): " Railroads and American Economic Growth: A "Market Access" Approach *," *Quarterly Journal of Economics*, 131, 799–858.

GENERAL ASSEMBLY OF RHODE ISLAND (1784): "Gradual Emancipation Act," Library of Congress: https://lccn.loc.gov/90898154.

GOLDSTEIN, J. R., M. ALEXANDER, C. BREEN, A. M. GONZÁLEZ, F. MENARES, M. OSBORNE, M. SNYDER, AND U. YILDIRIM (2021): "CenSoc Mortality File: Version 2.0. Berkeley: University of California, 2021." dataset: https://censoc.berkeley.edu/data/.

LEEUWEN, M. H. D. V. AND I. MAAS (2011): *HISCLASS: A Historical International Social Class Scheme*, Universitaire Pers Leuven.

LOGAN, T. D. AND J. M. PARMAN (2017): "The national rise in residential segregation," *Journal of Economic History*, 77, 127–170.

MANSON, S., J. SCHROEDER, D. V. RIPER, T. KUGLER, AND S. RUGGLES (2021): "IPUMS National Historical Geographic Information System: Version 16.0," dataset: http://doi.org/10.18128/D050.V16.0.

NATIONAL CENTER FOR HEALTH STATISTICS, U. (2021): "U.S. Life Expectancy at Birth by State and Census Tract - 2010-2015," dataset: https://data.cdc.gov/NCHS/U-S-Life-Expectancy-at-Birth-by-State-and-Census-T/5h56-n989.

NEW JERSEY STATE LEGISLATURE (1804): "An act for the gradual abolition of slavery," https://www.loc.gov/item/rbpe.0990100b/.

NEW YORK STATE LEGISLATURE (1799): "An act for the gradual abolition of slavery," New York laws of 1799, Chapter 62, p. 1: https://webarchive.loc.gov/all/20101110015236/http://iarchives.nysed.gov/dmsBlue/viewImageData.jsp?id=177879.

——— (1817): "An act relative to slaves and servants," New York State Archives: http://www.archives.nysed.gov/education/act-relative-slaves-and-servants-1817.

PENNSYLVANIA GENERAL ASSEMBLY (1780): "The Gradual Abolition Act," https://www.mountvernon.org/library/digitalhistory/digital-encyclopedia/article/gradual-abolition-act-of-1780/.

RUGGLES, S., S. FLOOD, R. GOEKEN, J. GROVER, E. MEYER, J. PACAS, AND M. SOBEK (2020): "IPUMS USA: Version 10.0," dataset: https://doi.org/10.18128/D010.V10.0.

SAAVEDRA, M. AND T. TWINAM (2020): "A machine learning approach to improving occupational income scores," *Explorations in Economic History*, 75, 101304.

SEGUIN, C. AND D. RIGBY (2019): "National crimes: A new national data set of lynchings in the United States, 1883 to 1941," *Socius*, 5, 2378023119841780.

THEIL, H. (1972): "Statistical decomposition analysis; with applications in the social and administrative sciences no v.14," in *Studies in Mathematical and Managerial Economics*, Amsterdam.

ZILVERSMIT, A. (1968): "Quok Walker, Mumbet, and the Abolition of Slavery in Massachusetts," *William and Mary Quarterly*, 25, 614–624.

B. Interpreting the Free-Enslaved Gap

Using the model in Section 5.1, the average difference between descendants of the Enslaved and the Free in 1940 is a consistent estimator of

$$\mathbb{E}[y_{i,1} \mid s_i = 1] - \mathbb{E}[y_{i,1} \mid s_i = 0] = (\lambda + \rho_{l(i,0)}) \left(\mathbb{E}[\alpha_{i,0} \mid s_i = 1] - \mathbb{E}[\alpha_{i,0} \mid s_i = 0]\right) + \mathbb{E}\left[\rho_{l(i,0)} \left(\gamma^0_{l(i,0)} - \delta\right) + \gamma^1_{l(i,1)} - \mid s_i = 1\right] - \mathbb{E}[\rho_{l(i,0)}\gamma^0_{l(i,0)} + \gamma^1_{l(i,1)} \mid s_i = 0]. \quad (11)$$

The difference between the two groups thus reflects 1) any potential differences in ability between the two groups, 2) different exposure to locations over time (as a result of slavery and of potential selection), and 3) the inherent disadvantage of descending from an enslaved person conditional on environment and ability. We think of the "effect of descending from enslaved people" as the expected difference between the two groups holding constant any confounding factors (i.e., ability). That is, we define the conditional treatment effect as

$$\begin{aligned} CTE(a_{i,0}) \equiv & \mathbb{E}[y_{i,1} \mid s_i = 1, \alpha_{i,0}] - \mathbb{E}[y_{i,1} \mid s_i = 0, \alpha_{i,0}] = \\ & = \mathbb{E}[\rho_{l(i,0)}\gamma^0_{l(i,0)} + \gamma^1_{l(i,1)} - \rho_{l(i,0)}\delta \mid s_i = 1, \alpha_{i,0}] - \mathbb{E}[\rho_{l(i,0)}\gamma^0_{l(i,0)} + \gamma^1_{l(i,1)} \mid s_i = 0, \alpha_{i,0}], \end{aligned} \quad (12)$$

and we define the average treatment effect as

$$ATE \equiv \int CTE(a_{i,0}) dF(\alpha_{i,0}). \quad (13)$$

From here, the Free-Enslaved gap in (11) is equal to

$$\mathbb{E}[y_{i,1} \mid s_i = 1] - \mathbb{E}[y_{i,1} \mid s_i = 0] = ATE - B, \quad (14)$$

where the (negative of) the selection bias B is

$$\begin{aligned} B = & \underbrace{\mathbb{E}[\left(\lambda + \rho_{l(i,0)}\right) \alpha_{i,0} \mid s_i = 0] - \mathbb{E}\left[\left(\lambda + \rho_{l(i,0)}\right) \alpha_{i,0} \mid s_i = 1\right]}_{\text{Potential selection into being free}} + \\ & \underbrace{\left(\mathbb{E}[\rho_{l(i,0)}\gamma^0_{l(i,0)} + \gamma^1_{l(i,1)} \mid s_i = 0] - \int \mathbb{E}[\rho_{l(i,0)}\gamma^0_{l(i,0)} + \gamma^1_{l(i,1)} \mid s_i = 0, \alpha_{i,0}] dF(\alpha_{i,0})\right)}_{\text{Potential selection into location by (descendants of) the Free}} - \\ & \underbrace{\left(\mathbb{E}\left[\rho_{l(i,0)} \left(\gamma^0_{l(i,0)} - \delta\right) + \gamma^1_{l(i,1)} \mid s_i = 1\right] - \int \mathbb{E}\left[\rho_{l(i,0)} \left(\gamma^0_{l(i,0)} - \delta\right) + \gamma^1_{l(i,1)} \mid s_i = 1, \alpha_{i,0}\right] dF(\alpha_{i,0})\right)}_{\text{Potential selection into location by (descendants of) the Enslaved}}. \end{aligned} \quad (15)$$

That is, the gaps between the two groups reflect the "treatment effect" of descending from an Enslaved ancestor but also potential differences in ability between the two groups that have persisted over time.

B.1 Selection into Being Free

Naturally, if being free before the Civil War was a matter of pure chance, the differences between the Free and the Enslaved have a causal interpretation.

Assumption A1. *Free-Enslaved status was independent of ability. That is,*

$$\alpha_{i,0} \perp\!\!\!\perp s_i.$$

That is, under Assumption A1, the bias in (15) is zero such that (14) shows that the unconditional gap is equal to the average treatment effect of being enslaved until the Civil War (1861–1865). A priori, Assumption A1 is strong. However, the plausibility of the assumption depends crucially on the conditions under which freedom was attained.

There were five main channels into freedom between the Revolutionary War (1775–1783) and the abolition of slavery in 1865: 1) by emancipation through abolition of slavery in the North in the late 18th and early 19th century, 2) by manumission through one's master, 3) by manumission through self-purchase, 4) by manumission through purchase by a third party, or 5) by running away. A person born to a free mother inherited their mother's freedom. In rare occasions, enslaved people were unintentionally freed by accompanying their masters on a trip to a free state. Setting foot on free soil freed enslaved people by law and some sued to enforce their rights (see, e.g., Rose, 2009).

In 1860, around half of the free population was born in the North, which we argue is a reasonable approximation of the share of the free families who were freed through general emancipation in the North. Within the remaining half, it is hard to estimate the share of people who were freed "legally" and those who ran away. While the 1850 and 1860 censuses suggest the number of runaway slaves is less than 2,000 (out of a total population of around 4 million), the true number is likely much higher (Franklin and Schweninger, 2000).

Dittmar and Naidu (2012) use runaway slave advertisements placed in Southern newspapers between 1840 and 1860 and suggest that such advertisements were placed for around 8,000 runaway slaves throughout those two decades. However, the authors also point out that "it is clear that among the many absconders only a small fraction remained at large for a lengthy period." The odds of a successful escape were especially small in the Lower South. This is corroborated by the fact that in a Pennsylvania census of Free Black Americans, only 2 out of 314 people who were not born free indicated that they at-

tained freedom through escape.[34] It is therefore safe to conclude that the vast majority of those who became free in the South did so through manumission (as opposed to escape).

Since slavery had been de facto abolished in the North by 1850 (see Table A.2), the enslaved people there were freed non-selectively. That is, as long as one is willing to assume that those enslaved in the North were not inherently different from those enslaved in the South, those in the North were freed entirely independent of any observed or unobserved characteristics. In the South, the degree of selection into manumission varied largely across time and locations. Around the 1780s, the early years after the Revolutionary War, there was a stream of manumissions motivated by morality or religion. In later antebellum years, manumission turned into an instrument to uphold slavery (Berlin, 1974). It did not, in most cases, arise from anti-slavery sentiments. On the contrary, many owners manumitted their slaves as a reward for loyalty and by doing so "reinforced rather than challenged the values, assumptions, and discipline of slavery" (Wolf, 2006, p. 44).

One could imagine that the practice of manumission induced a degree of selection into being free. Indeed, some quantitative evidence on the presence of selection into manumission exists. Cole (2005) finds that in Louisiana, manumitted people were 62.5 percent female (43.6 percent in the enslaved population) and much more likely to be "mulatto" (38.5 percent) than the slave population (5.8 percent). This is consistent with the observation that manumission in the Lower South was reserved for "illicit offspring, special favorites, or least productive slaves" (Berlin, 1974). Bodenhorn (2011), too, finds evidence of preferential manumission for people of mixed race in Virginia. Similarly, Berlin (1974) argues that skilled slaves had a larger chance of accumulating enough wealth to be manumitted through self-purchase. Little is known about selection into being manumitted through purchase by other people (usually other free Black people). Runaways, however, "as a group, had always been more skilled, sophisticated, and aggressive than the mass of slaves" (Berlin, 1974, p. 160). Table B.3 summarizes the discussion.

To assess the potential bias that may result from selection into freedom on observables, we re-estimate the Free-Enslaved gap using only the Free who had not accumulated any measurable physical or human capital by the end of slavery. Even for this subsample, we find large Free-Enslaved gaps in 1940 (see Appendix Table J.13). This result provides additional evidence against there being substantial bias in the Free-Enslaved gap through selection into freedom.

B.2 Bounding the Selection Bias

The assumption of independence of Free-Enslaved status and ability (Assumption A1) may be strong. In this section, we show that under weaker assumptions, we can provide

[34] Pennsylvania Abolition Society and Society of Friends Manuscript Census Schedules, 1838. Available in machine-readable form through https://doi.org/10.3886/ICPSR03805.v1.

TABLE B.3: Relative prevalence of and selectivity in different roads to freedom

	%	Degree of selection
Emancipation in North	≈ 50	None
Manumission by master	30-40	Varied across time and locations
Manumission by self-purchase	5-10	Potentially high
Manumission by a third buyer	5-10	Unknown
Escape	< 5	Potentially high

Notes: This table indicates a rough breakdown of the relative probability of attaining freedom in various ways. The percentage emancipated in the North is estimated by the fraction of free Black people born in the North in the 1860 census. The fraction that escaped is a conservative upper bound given the observations mentioned in the text. The remaining probability is attributed to manumissions. The distribution within manumissions is derived from (Bodenhorn, 2011): 10-20 percent through self-purchase, 10-20 percent through a third buyer, and the remaining 60-80 percent by the master.

plausible bounds for the average treatment effect defined in (13). Assumptions B1 and B2 are the key assumptions.

Assumption B1. *The initial (pre-1865) distribution across states of the Free and the Enslaved was not a result of any potential selection into being free. That is,*

$$\mathbb{E}\left[\rho_{l(i,0)}\left(\gamma^0_{l(i,0)} - \delta s_i\right) \mid s_i = s\right] = \int \mathbb{E}\left[\rho_{l(i,0)}\left(\gamma^0_{l(i,0)} - \delta s_i\right) \mid s_i = s, \alpha_{i,0}\right] dF(\alpha_{i,0}) \quad \forall s \in \{0,1\}.$$

Assumption B1 requires that, in the presence of any selection into being free, selection did not affect the overall geographic distribution of either the Free or Enslaved. Note that Assumption A1 is sufficient but not necessary for Assumption B1 to be satisfied. To make the assumption as plausible as possible, we let $l(i,0)$ be the state of birth in 1870 (as opposed to the state of residence). For the formerly enslaved population, this thus refers to the state where they were enslaved at their time of birth. In Section 6, we argue in detail that selection into enslavement location was likely limited, mainly because the enslaved population did not enjoy freedom of movement (this condition is sufficient but not necessary for Assumption B1 to hold). The free population in 1870 consisted both of people who were born free and those who were born enslaved and freed at some point in their life before 1860. For the latter group, $l(i,0)$ also refers to the location of enslavement.

The remaining threat to Assumption B1 is that those who were born free may have had parents who selected themselves into certain locations in the first half of the 19th century. Around half of the freeborn descended from people living in the North who were (non-selectively) freed by general emancipation in the Northern states. One may be concerned, however, that selection arose through the free population in the South. Their geographic distribution could have differed from the counterfactual geographic distribution of the Free under no selection into being free. We argue this is unlikely to have made a meaningful difference. Migration was limited among the Free, and therefore their state

of residence was strongly influenced by their family's location during enslavement. For some of the Free, this was simply due to a lack of migration opportunities (Morgan, 2012, p. 490). However, for many others, there was also a strong unwillingness to migrate. A contemporaneous observer in Charleston, South Carolina wrote that "the Free People [of color] will never emigrate; they have so little to hope for, and so much to dread, from any change of place, that they will adhere to the spot of their nativity, under the pressure of any inconvenience."[35]

Furthermore, even though political and civil liberties for Black people were greater in the North, economic opportunities tended to be considerably larger in the South, which "suggests why so few southern freemen sought to leave the slave states for 'freedom' in the North or, for that matter, why so few northern blacks went South in search of better jobs. While southern free Negroes could expand their political and social liberties greatly by migrating to a free state, they could do so only at the expense of their economic opportunities (Berlin, 1976, p. 313)." All in all, this meant that in the antebellum period there was no clear "destination" in which free people could select themselves. Therefore, we argue that it is implausible that selection into being free affected the geographic distribution of the Free by much.

Assumption B2. *Selection into freedom in the South was, if anything, stronger than overall selection into freedom. That is,*

$$\sum_{l\in\mathcal{L}}\left(\lambda+\rho_{l(i,0)}\right)\left(\mathbb{E}\left[\alpha_{i,0}\mid s_i=0,l(i,0)=l\right]-\mathbb{E}\left[\alpha_{i,0}\mid s_i=1,l(i,0)=l\right]\right)w(l)\geq \\ \mathbb{E}\left[\left(\lambda+\rho_{l(i,0)}\right)\alpha_{i,0}\mid s_i=0\right]-\mathbb{E}\left[\left(\lambda+\rho_{l(i,0)}\right)\alpha_{i,0}\mid s_i=1\right],$$

where

$$w(l)=\frac{\Pr\left(s_i=1\mid l(i,0)=l\right)\Pr\left(s_i=0\mid l(i,0)=l\right)\Pr\left(l(i,0)=l\right)}{\sum_{l\in\mathcal{L}}\Pr\left(s_i=1\mid l(i,0)=l\right)\Pr\left(s_i=0\mid l(i,0)=l\right)\Pr\left(l(i,0)=l\right)} \tag{16}$$

is the least squares regression weight for location l.

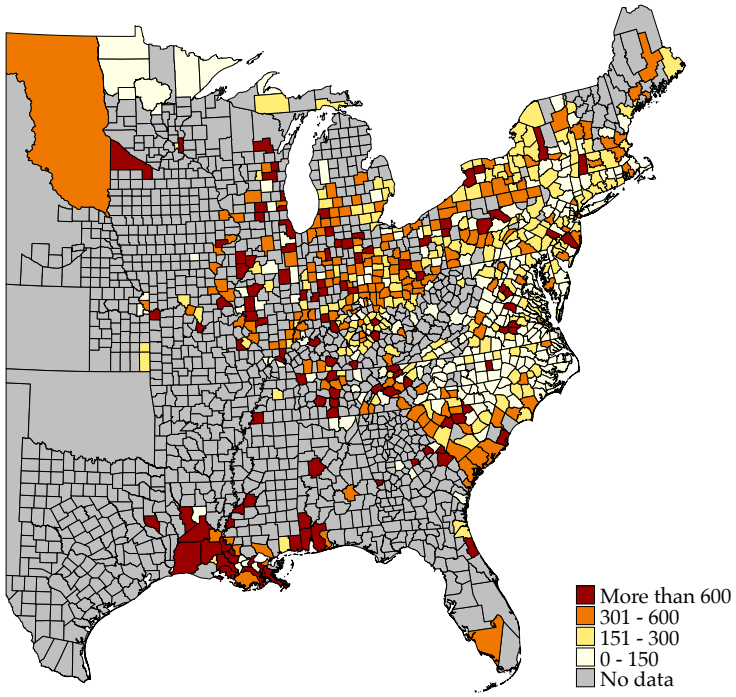
Note that $w(l)$ is positive only for locations where both free and enslaved people lived before Emancipation. It is thus zero for locations in the North and is largest in the Upper South where there were substantial populations of both enslaved and free people. Since the Free in the North mostly achieved freedom through general emancipation, they would likely have been less selected than the Free in the South. Also, if there was any "negative selection" into being enslaved up to 1865, this would likely have been stronger in the Upper South (where more people were free) than in the Lower South. This is corroborated by the fact that free Black Americans in the Upper South tended to have a far lower socioeconomic status than those in the Lower South (see Appendix Figure B.4).

[35] Quoted in Berlin (1974, p. 146).

Last, if anything, one would expect post-slavery relative mobility to have been lower in the South (i.e., a higher ρ_l). All observations point to Assumption B2 being satisfied.

To bound the bias under Assumptions B1 and B2, we need to impose two more assumptions that we argue are almost trivial.

FIGURE B.4: Free Black Americans' Average Wealth in 1860



Notes: This figure shows the average total wealth (real and personal property) among free Black Americans in the 1860 census by county. The maps are truncated to omit the western half of the country, which at the time was only sparsely populated.

Assumption B3. *In the absence of selection into being free before the Civil War, descendants of the free population would have been able to locate themselves at least as favorably as descendants of the Enslaved conditional on initial location. That is,*

$$\sum_{l\in\mathcal{L}} \int \mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 0, l(i,0) = l, a_{i,0}] dF(a_{i,0}) w(l) \geq \sum_{l\in\mathcal{L}} \int \mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 1, l(i,0) = l, a_{i,0}] dF(a_{i,0}) w(l). \quad (17)$$

Assumption B3 rules out the possibility that being enslaved had a *positive* causal effect on the 1940 location given ancestor location. The Free tended to have more resources and networks available to migrate to places of opportunity. Indeed, Appendix Figure D.7 shows that the Free were considerably more likely to move from the South to the North between 1870 and 1940 (even when controlling for ancestor county).

Assumption B4. *Had the Free been endowed with the same geographical distribution as the Enslaved, the differential selection into their location in 1940 would have, if anything, been stronger. That is,*

$$\sum_{l\in\mathcal{L}}\left(\mathbb{E}[\alpha_{l(i,1)} \mid s_i=0, l(i,0)=l] - \int \mathbb{E}[\alpha_{l(i,1)} \mid s_i=0, l(i,0)=l, a_{i,0}]dF(a_{i,0})\right) w(l) -$$
$$\sum_{l\in\mathcal{L}}\left(\mathbb{E}[\alpha_{l(i,1)} \mid s_i=1, l(i,0)=l] - \int \mathbb{E}[\alpha_{l(i,1)} \mid s_i=1, l(i,0)=l, a_{i,0}]dF(a_{i,0})\right) w(l)$$
$$\geq \left(\mathbb{E}[\alpha_{l(i,1)} \mid s_i=0] - \int \mathbb{E}[\alpha_{l(i,1)} \mid s_i=0, a_{i,0}]dF(a_{i,0})\right) -$$
$$\left(\mathbb{E}[\alpha_{l(i,1)} \mid s_i=1] - \int \mathbb{E}[\alpha_{l(i,1)} \mid s_i=1, a_{i,0}]dF(a_{i,0})\right)$$

Loosely speaking, Assumption B4 requires that the differential selection into location would have been larger had the geographical distribution of the Free and the Enslaved been the same, entirely in the South, and predominantly in the Lower South.

We argue that Assumption B4 is very likely to hold. The potential for selection into location is larger whenever the current location is less beneficial. For instance, for the Free who were concentrated in the North and Upper South, there was relatively little room for selective migration since the Great Migration showed (by revealed preference) that they were already in the most favorable locations for Black people at the time. On the other hand, for the enslaved population in the South (especially the Lower South), the potential for an effect of selective migration was large since those locations tended to provide the least opportunities for progress. Had they been concentrated in the Upper South, this potential would have likely been smaller, if anything. Therefore, under this counterfactual geographical distribution, it seems plausible that selective migration would have been *more* important for the Free, while it would have been *less* important for the enslaved population. Both tend to the direction of satisfying Assumption B4.

Proposition 1. *Under Assumptions B1, B2, B3, and B4, the Free-Enslaved gap in 1940 conditional on ancestor location defines a bound on the sum of the (negative of the) selection bias and the direct effect of descending from enslaved ancestors in* (15) *from below. Formally,*

$$\sum_{l\in\mathcal{L}}\left(\mathbb{E}[y_{i,1} \mid s_i=0, l(i,0)=l] - \mathbb{E}[y_{i,1} \mid s_i=1, l(i,0)=l]\right) w(l) \geq B + \rho_{l(i,0)}\delta. \tag{18}$$

Proof: Using Assumption B2, the Free-Enslaved gap conditional on location at $t=0$ can

be bound by

$$\sum_{l\in\mathcal{L}} \left(\mathbb{E}[y_{i,1} \mid s_i = 0, l(i,0) = l] - \mathbb{E}[y_{i,1} \mid s_i = 1, l(i,0) = l]\right) w(l) \geq$$
$$\mathbb{E}\left[\left(\lambda + \rho_{l(i,0)}\right)\alpha_{i,0} \mid s_i = 0\right] - \mathbb{E}\left[\left(\lambda + \rho_{l(i,0)}\right)\alpha_{i,0} \mid s_i = 1\right] +$$
$$\sum_{l\in\mathcal{L}} \left(\mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 0, l(i,0) = l] - \mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 1, l(i,0) = l]\right) w(l) + \rho_{l(i,0)}\delta. \quad (19)$$

Assumptions B3 and B4 ensure we can bound this further by

$$\sum_{l\in\mathcal{L}} \left(\mathbb{E}[y_{i,1} \mid s_i = 0, l(i,0) = l] - \mathbb{E}[y_{i,1} \mid s_i = 1, l(i,0) = l]\right) w(l) \geq$$
$$\mathbb{E}\left[\left(\lambda + \rho_{l(i,0)}\right)\alpha_{i,0} \mid s_i = 0\right] - \mathbb{E}\left[\left(\lambda + \rho_{l(i,0)}\right)\alpha_{i,0} \mid s_i = 1\right] +$$
$$\left(\mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 0] - \int \mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 0, a_{i,0}]dF(a_{i,0})\right) -$$
$$\left(\mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 1] - \int \mathbb{E}[\gamma^1_{l(i,1)} \mid s_i = 1, a_{i,0}]dF(a_{i,0})\right) + \rho_{l(i,0)}\delta = B + \rho_{l(i,0)}\delta, \quad (20)$$

where the last equality follows from (15) and Assumption B1. □

The left-hand side in (18) is exactly the estimand in an ordinary least squares regression of the outcome $y_{i,1}$ on ancestor slave status s_i and ancestor location $l(i,0)$ (Angrist and Pischke, 2008).

REFERENCES

ANGRIST, J. D. AND J.-S. PISCHKE (2008): *Mostly harmless econometrics*, Princeton university press.

BERLIN, I. (1974): *Slaves without masters: The free Negro in the antebellum South*, Oxford University Press.

——— (1976): "The Structure of the Free Negro Caste in the Antebellum United States," *Journal of Social History*, 9, 297–318.

BODENHORN, H. (2011): "Manumission in nineteenth-century Virginia," *Cliometrica*, 5, 145–164.

COLE, S. (2005): "Capitalism and freedom: Manumissions and the slave market in Louisiana, 1725–1820," *Journal of Economic History*, 65, 1008–1027.

DITTMAR, J. AND S. NAIDU (2012): "Contested Property: Fugitive Slaves in the Antebellum U.S. South," Tech. rep.

FRANKLIN, J. H. AND L. SCHWENINGER (2000): *Runaway slaves: Rebels on the plantation*, OUP USA.

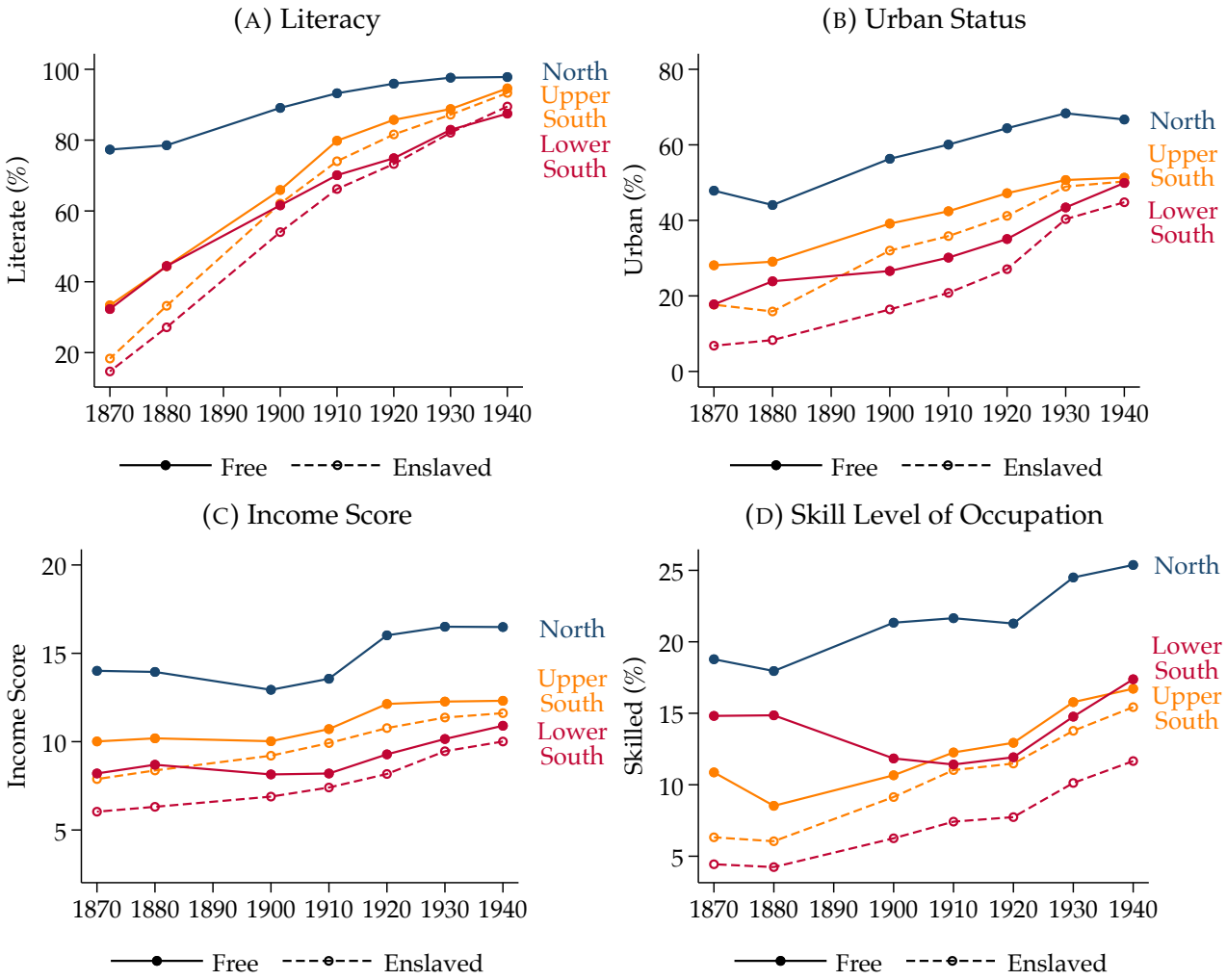
MORGAN, P. D. (2012): *Slave counterpoint: Black culture in the eighteenth-century Chesapeake and Lowcountry*, UNC Press Books.

ROSE, B. Z. (2009): *Mother of Freedom: Mum Bett and the Roots of Abolition*, TreeLine Press.

WOLF, E. S. (2006): *Race and Liberty in the New Nation: Emancipation in Virginia from the Revolution to Nat Turner's Rebellion*, LSU Press.

C. Socioeconomic Status among Descendants of the Free and the Enslaved

FIGURE C.5: Socioeconomic Characteristics of Family by Region of Origin (1870-1940)

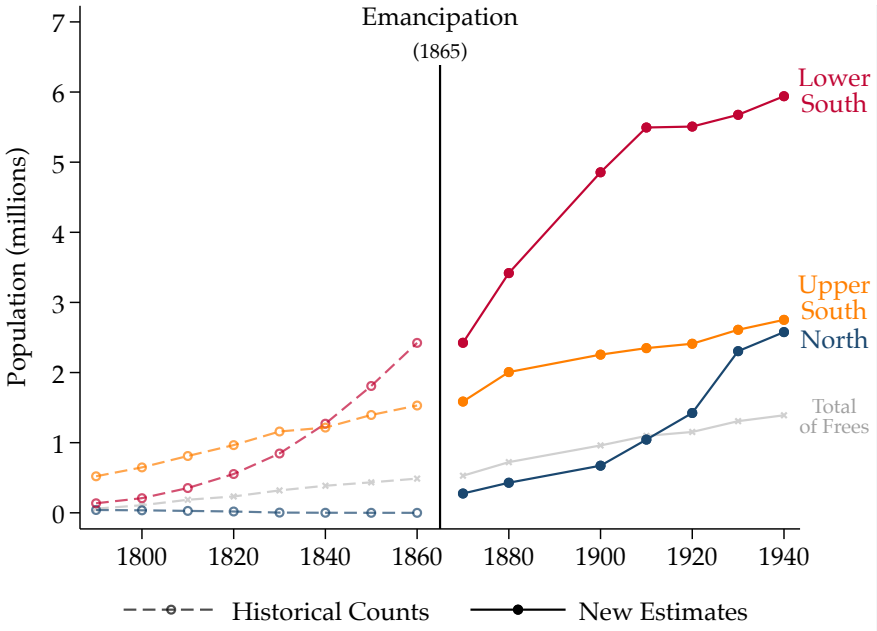


Notes: This figure shows the averages of characteristics in the cross-section of prime-age male descendants of the Free and the Enslaved by their ancestor's region (family's residence pre-1880). Incomes Score uses the Lido score developed by Saavedra and Twinam (2020). Skill level of occupations is inferred from the classification by Leeuwen and Maas (2011). See Data Appendix A for details on the sample and data.

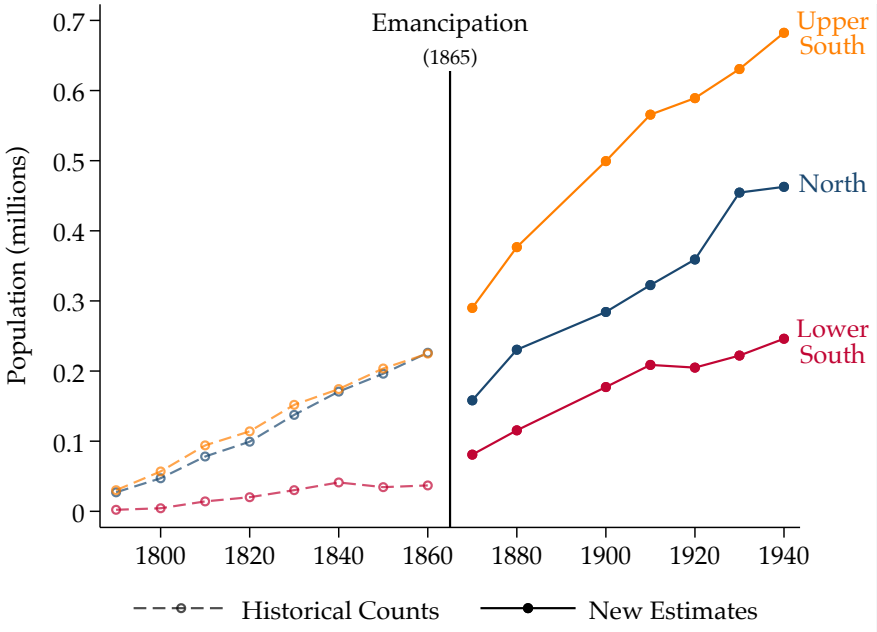
D. Migration among Descendants of the Free and the Enslaved

Figure D.6: Population of the Free and the Enslaved by Region (1790-1940)

(A) The Enslaved and Their Descendants

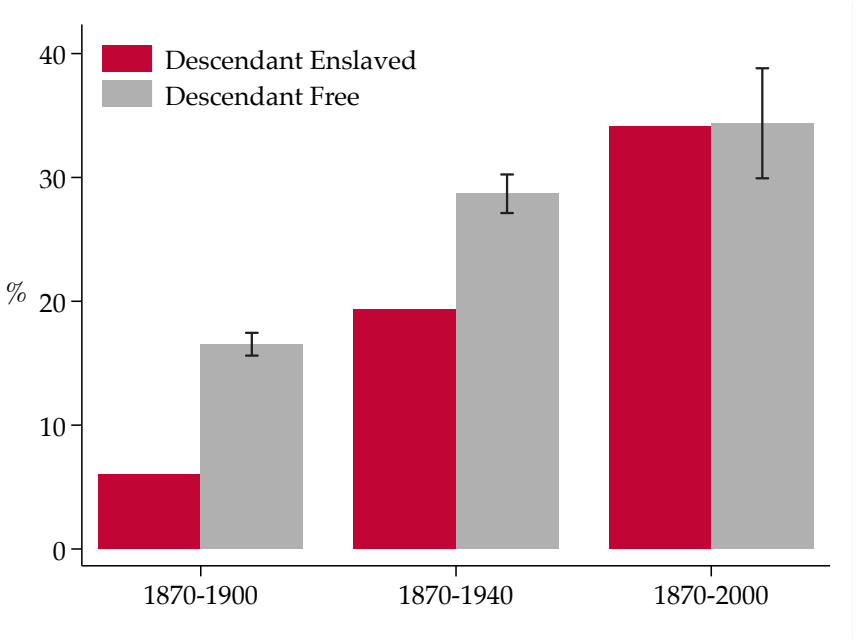


(B) The Free and Their Descendants



Notes: This figure shows the number of enslaved (Panel A) and free (Panel B) Black Americans and their descendants across regions. Dashed lines represent aggregate counts as reported in historical records (Berlin, 1974). Solid lines represent population counts inferred from our data. We compute the fraction of the Free and the Enslaved living in each region and multiply by the total number of Black Americans who descend from the Free and the Enslaved respectively. This procedure assumes that both population grow at a similar rate, implying that each accounts for a fraction of the Black population that is constant at the 1860 level.

FIGURE D.7: Probability of Participating in the Great Migration



Notes: This figure shows the fraction of descendants of the Free and the Enslaved who live in the North in 1900 (before the Great Migration), 1940 (after the first wave of the Great Migration), and 2000 (around 30 years after the end of the Great Migration), conditional on their family residing in the South in 1870. The error bar indicates the 95 percent confidence interval for the difference between the two groups. Descendants of free Black Americans are 50 percent more likely to participate in the Great Migration until 1940 than descendants of the Enslaved. The gap remains constant even after controlling for their socioeconomic status. Controlling for their family's county of residence in 1870, however, shrinks this gap by half. The geographic location of enslaved ancestors substantially delayed the participation in the Great Migration of their descendants, potentially also delaying their economic progress after slavery.

FIGURE D.8: Black Families Leaving the Slave States by 1870 State of Origin

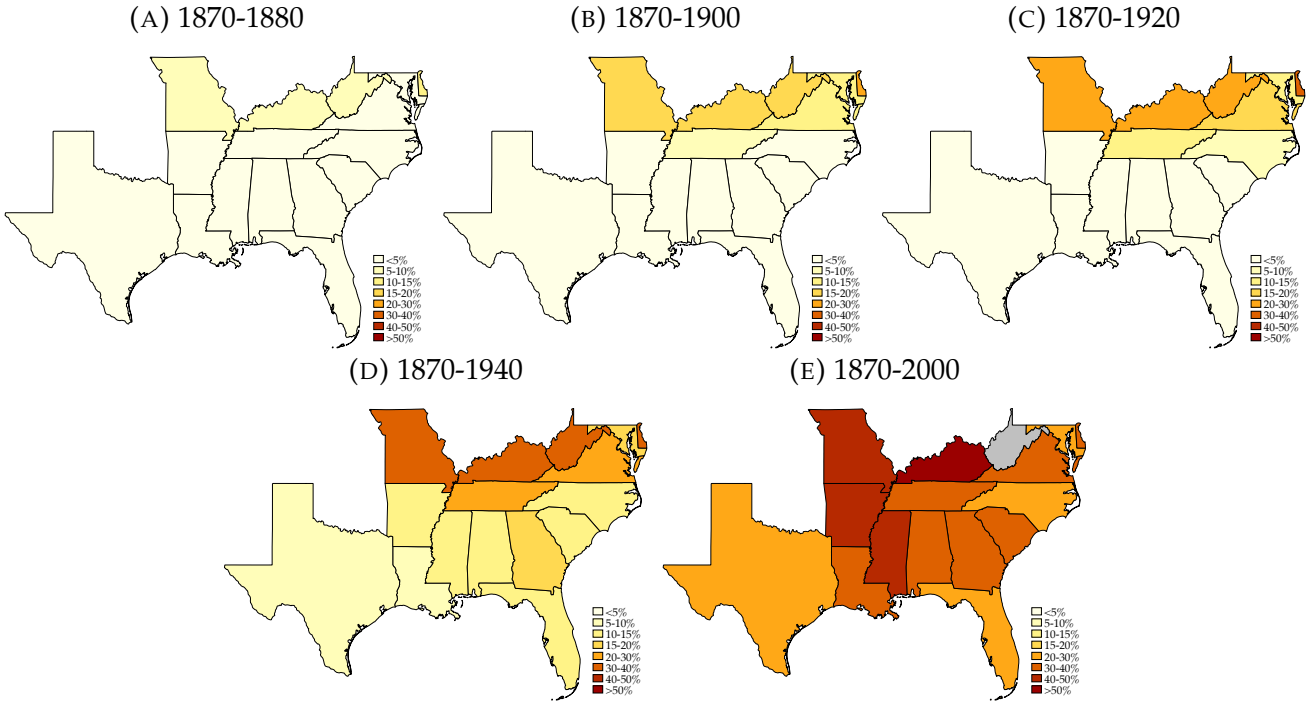
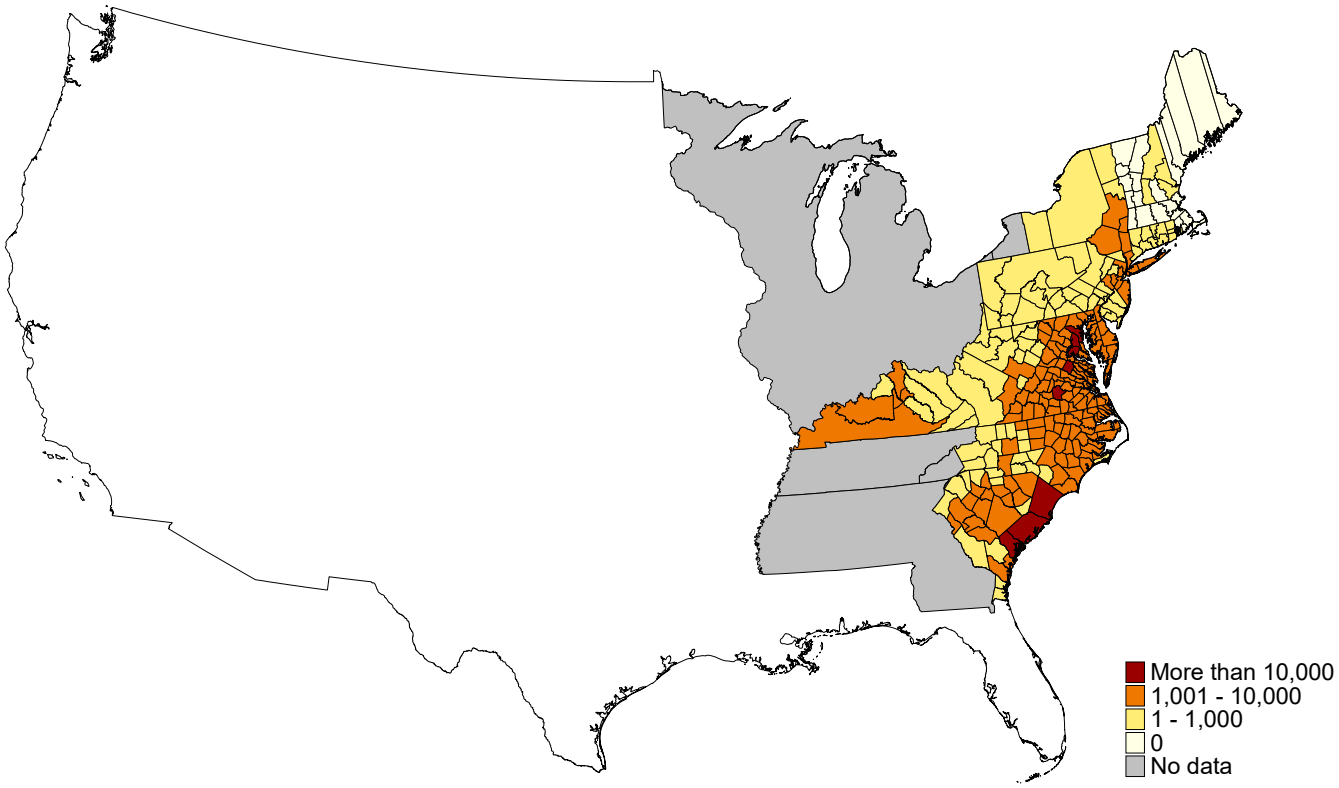
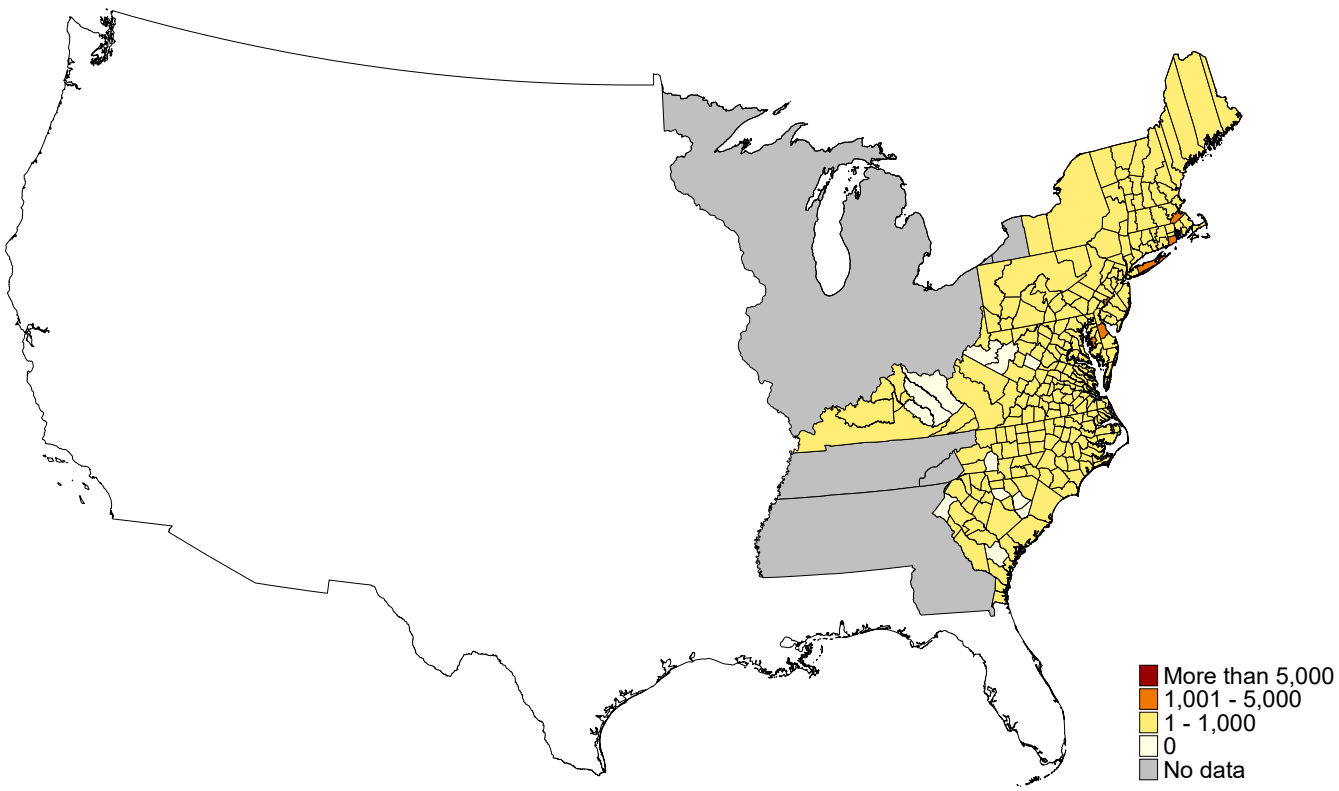


FIGURE D.9: County Population of Enslaved and Free (1790 & 1860)

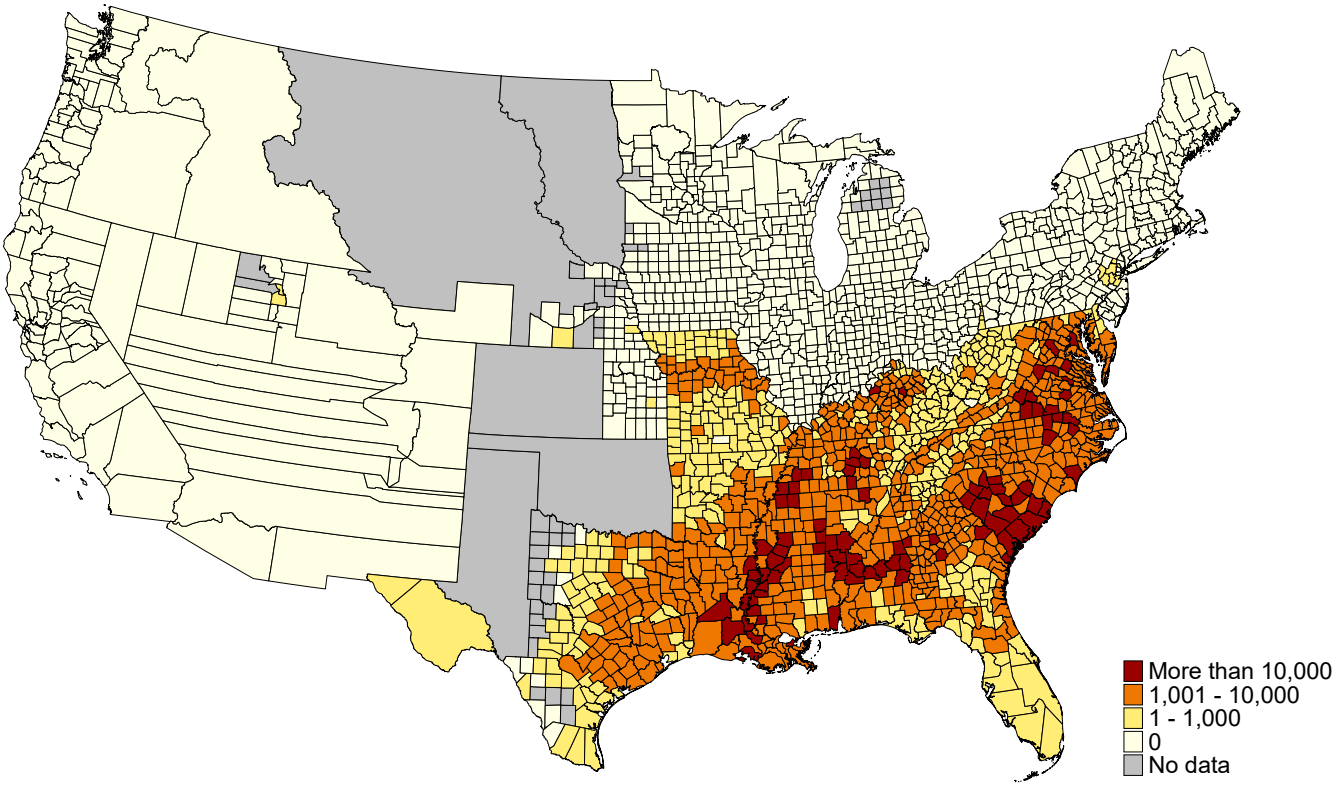
(A) The Enslaved in 1790



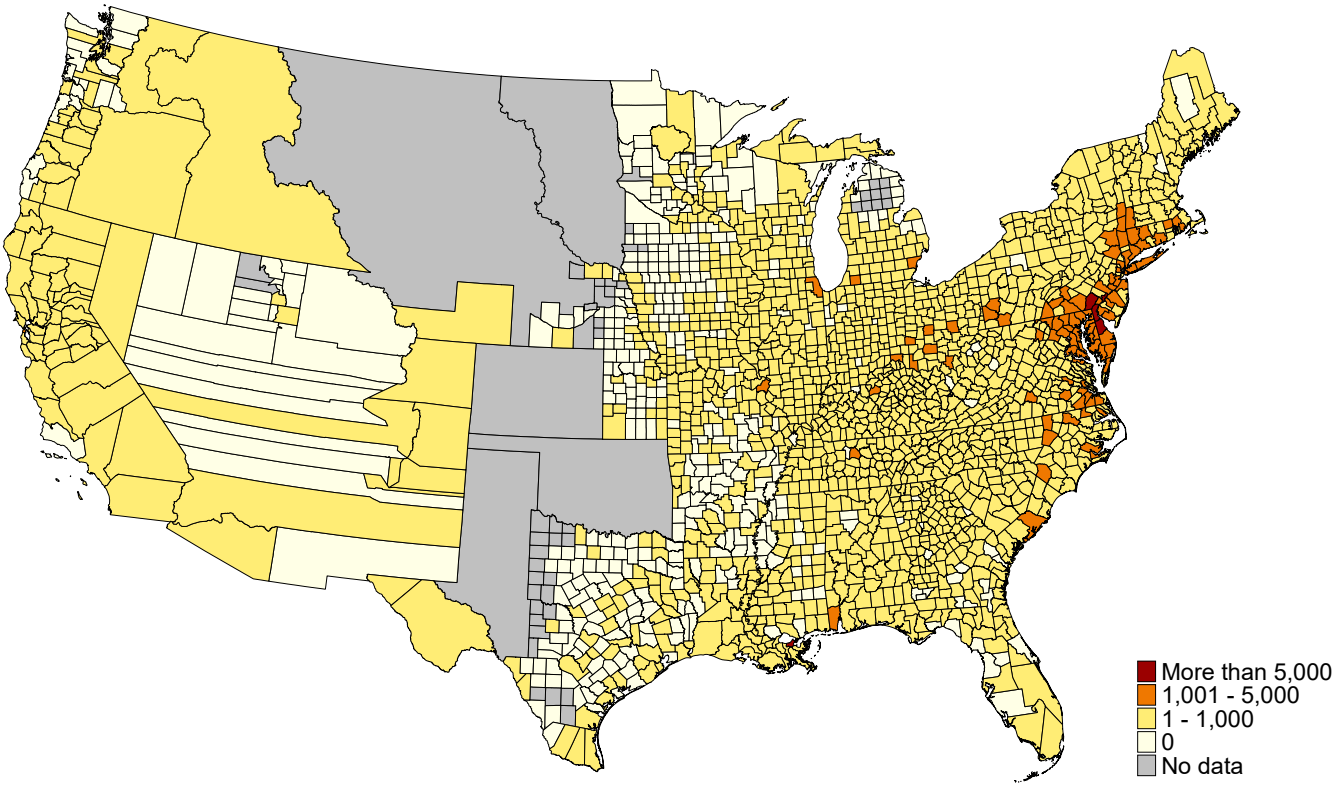
(B) The Free in 1790



(C) The Enslaved in 1860



(D) The Free in 1860



E. Adjusting Estimates for Misclassification Bias

We may incorrectly classify a substantial number of Black families as descendants of the Enslaved if they are born in a slave state where a large share of Black Americans was free before 1860. For example, among Black Americans from Maryland, around 50 percent are free in 1860. Due to imperfect linking rates, many of those free Black families will be misclassified as Enslaved by our method.

We adjust our estimates for bias that may arise from this type of misclassification. We make use that our original estimates are a weighted average of the (unknown) unbiased estimate and the non-causal estimate for free Black Americans:

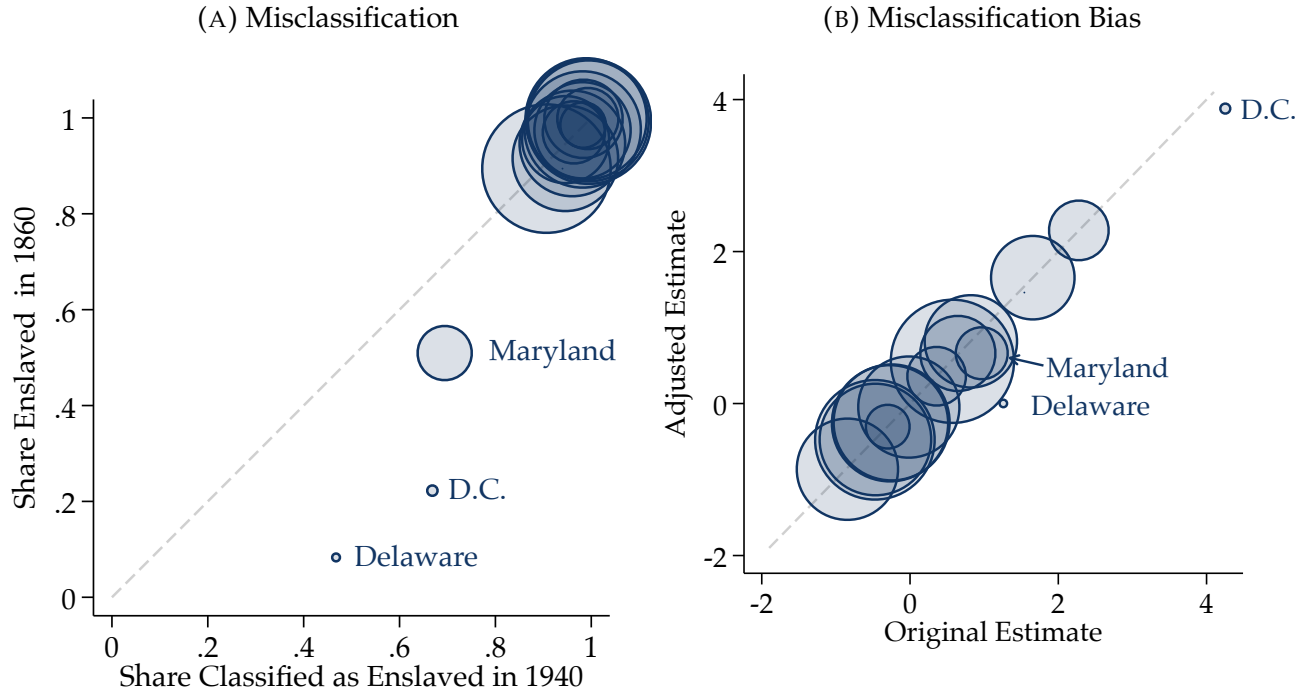
$$\widehat{\beta}_{\text{original}} = \frac{\text{Enslaved}_{s,\text{links}}}{\text{Enslaved}_{s,1860}} \cdot \widehat{\beta}_{\text{unbiased}} + \left(1 - \frac{\text{Enslaved}_{s,\text{links}}}{\text{Enslaved}_{s,1860}}\right) \cdot \widehat{\beta}_{\text{free}}, \tag{21}$$

where $\text{Enslaved}_{s,\text{links}}$ is the share of Black Americans who descend from the Enslaved of state s according to our classification in 1940, $\text{Enslaved}_{s,1860}$ is the true share of Black Americans who descend from the Enslaved of state s according to the 1860 census, and $\widehat{\beta}_{\text{free}}$ is the non-causal estimate for outcomes of free Black families with ancestors from state s.

In our sample, for instance, 70 percent of Black Americans with ancestors from Maryland are classified as descendants of the Enslaved in 1940—almost 20 percentage points more than the true share. Comparing the two groups suggests, unsurprisingly, that the free families of Maryland achieved somewhat higher levels of education by 1940 than descendants of the Enslaved. As a consequence, our original estimates (that incorrectly include some free families) are upward biased estimates of the true effect living in Maryland had on descendants of the Enslaved. Using the equation above, we can adjust our estimates.

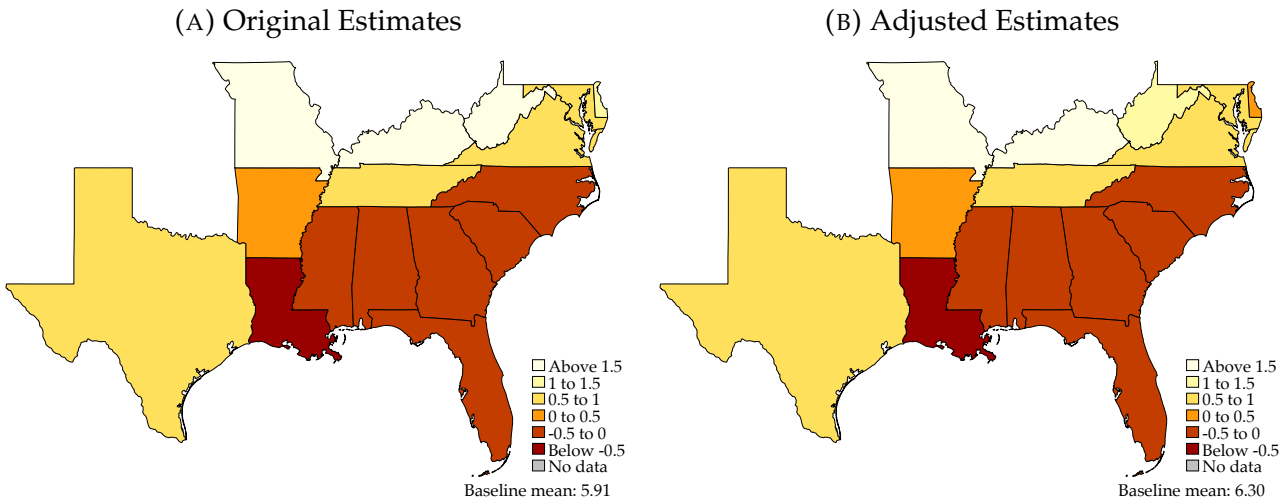
Adjusting our estimates for bias that may result from this form of misclassification turns out not to quantitatively matter. Figure E.10 shows that the share of Black Americans that descend from the Enslaved only deviates substantially from our classification for three small slave states. Accordingly, adjusting our original estimates of the causal effect of each state barely affects our estimates. Panel (A) of Figure E.11 shows that the original estimates tend to almost perfectly coincide with the adjusted estimates. Panel (B) shows the map of states' effects before and after the adjustment.

FIGURE E.10: Misclassification and Bias



Notes: This figure assesses on misclassification of the Free-Enslaved status and the impact misclassification has on our estimates. Panel (A) shows the extent of misclassification as descendants of the Enslaved or the Free among Black Americans in 1940 with ancestors born in a given state before 1870. Panel (B) shows our causal estimates of living in each state before and after adjusting for misclassification bias. See Data Appendix A for details on the sample and data.

FIGURE E.11: Original and Adjusted Estimates



Notes: This figure compares the geographic patterns in our original estimates of states' effect on Black economic progress after slavery with the adjusted estimates. See Data Appendix A for details on the sample and data.

F. ADJUSTING ESTIMATES FOR INTERMARRIAGE

Our empirical strategy identifies the gap between Black individuals whose ancestors along the male line were enslaved until the Civil War and those whose ancestors along the male line were free before the Civil War. Women tended to change their last names upon marriage, making it

impossible to assess the Free-Enslaved status of female ancestors. Therefore, some individuals who are classified as descendants of one group, may in truth also partly descend from the other group. This type of misclassification biases our estimate of the Free-Enslaved gap towards zero.

We use information on a person's birth place, their parents' birth places, and their spouse's birth place to assess the degree of intermarriage. In particular, we can assess the degree to which Black Americans from slave states and free states intermarried to proxy for how common intermarriage was. Note that the Free and the Enslaved were distributed very differently across America.

FIGURE F.12: Black Intermarriages between Descendants of the Free and the Enslaved

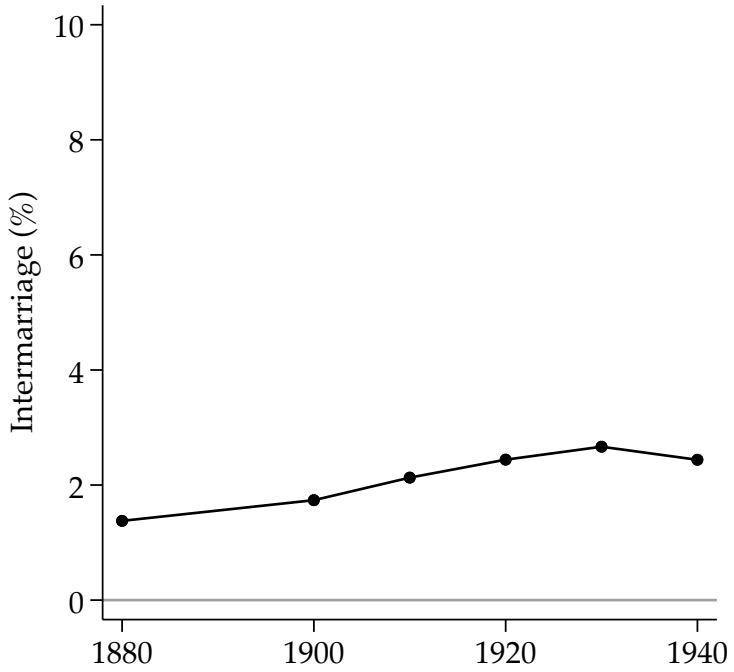


Figure F.12 shows the fraction of Black marriages that constituted intermarriage between descendants of the Free (mother born in free state prior to 1865) and the Enslaved (mother born in slave state prior to 1865). By 1940, less than 2 percent of all Black marriages constituted such intermarriage. Therefore, those who we classify as descendants of the Enslaved likely have no free ancestors. However, because the number of the Free is small, even this small degree of intermarriage makes it likely that some fraction of the people we classify as Free in 1940 have some ancestor who was enslaved.

We can use the aggregate proxy for intermarriage to correct the original estimate of the Free-Enslaved gap as follows:

$$\begin{aligned} \widehat{\beta}_t^{\text{true}} &= \overline{Y}_t^{\text{Ensl}} - \overline{Y}_t^{\text{Free}} \\ \widehat{\beta}_t^{\text{original}} &= \left(s_{t,\text{Ensl}}^{\text{IM}} \cdot \overline{Y}_t^{\text{Ensl}} + [1 - s_{t,\text{Ensl}}^{\text{IM}}] \cdot \overline{Y}_t^{\text{Free}}\right) - \left(s_{t,\text{Free}}^{\text{IM}} \cdot \overline{Y}_t^{\text{Free}} + [1 - s_{t,\text{Free}}^{\text{IM}}] \cdot \overline{Y}_t^{\text{Ensl}}\right) \end{aligned}$$

where $\widehat{\beta}_t^{\text{true}}$ is the (unknown) gap between a person who descends only from free Black Americans and a person who only descends from enslaved people, $\overline{Y}_t^{\text{Ensl}}$ and $\overline{Y}_t^{\text{Free}}$ are the average outcomes of descendants of the Enslaved and Free, and $s_{t,\text{Ensl}}^{\text{IM}}$ and $s_{t,\text{Free}}^{\text{IM}}$ are the share of En-

slaved and Free who engage in intermarriage. Using our estimates for 1940, we have

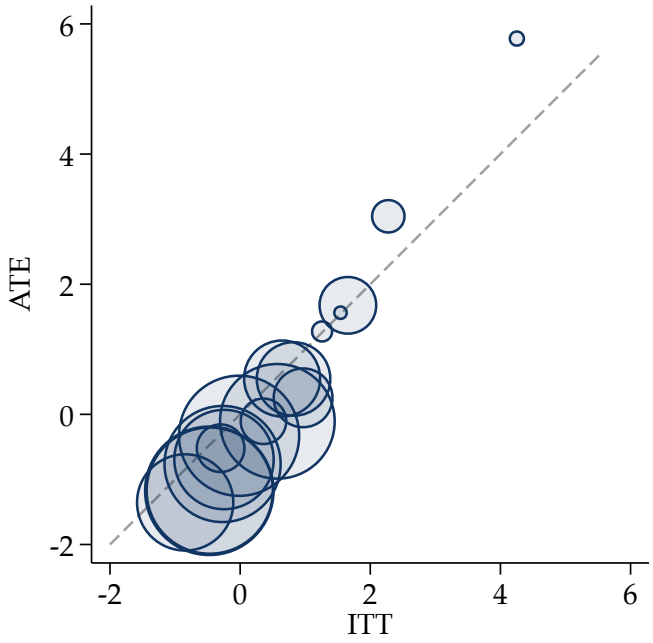
$$\begin{aligned}\widehat{\beta}_{1940}^{\text{original}} &\approx \left(0.01 \cdot \overline{Y}_t^{\text{Ensl}} + 0.99 \cdot \overline{Y}_t^{\text{Free}}\right) - \left(0.33 \cdot \overline{Y}_t^{\text{Free}} + 0.67 \cdot \overline{Y}_t^{\text{Ensl}}\right) \\ &\approx 0.66 \cdot \left(\overline{Y}_t^{\text{Free}} - \overline{Y}_t^{\text{Ensl}}\right)\end{aligned}$$

Thus, the true Free-Enslaved gap is around one-and-a-half times as large in 1940 as our original estimates would suggest. For earlier decades, the adjustment factor is somewhat smaller.

G. The Direct Effect of Locations After Accounting for Migration

Our estimates of how being freed in a given location affected the economic progress of Black families reflects both the effect the original location and the expected effects of future locations conditional on the 1870 location. We decompose this intent-to-treat (ITT) effect into (1) the direct effect of the location of origin and (2) the potential indirect effect of destinations that Black families may migrate to. To do so, we invert the location-to-location migration probability matrix to back out the ATE of living in each state from the estimated ITT effect.

FIGURE G.13: ITT Effect and ATE of Living in Each Southern State (1870–1940) on Years of Education in 1940



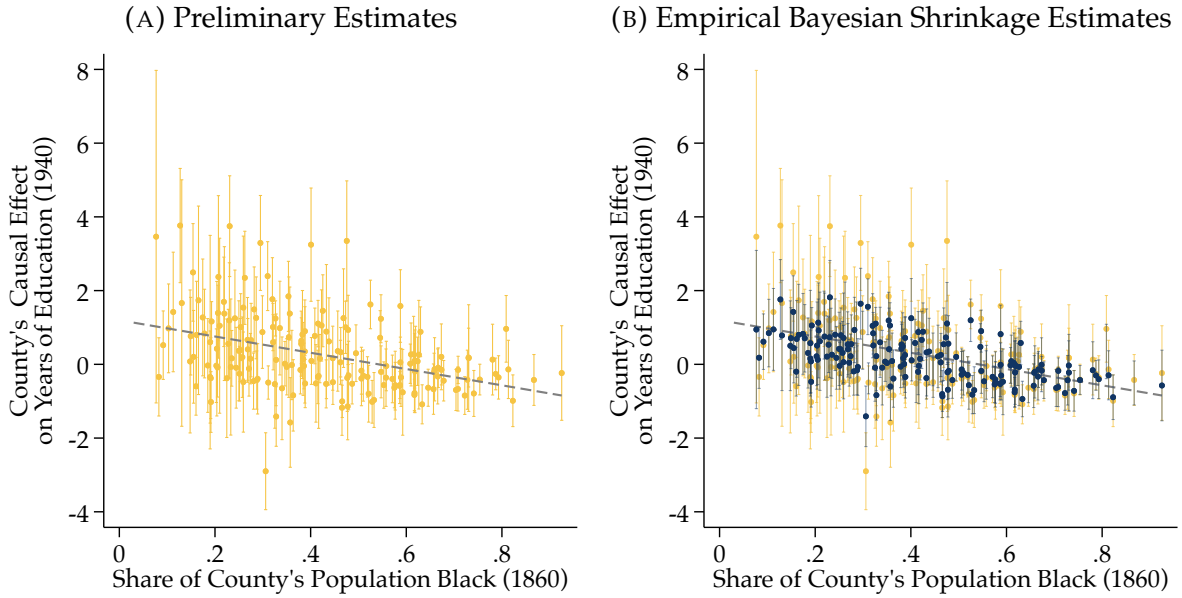
Notes: This figure compares our original (ITT) estimates of how being freed in a given state affected a Black family's economic progress to the direct (ATE) effect that living in that state had. The estimates are in years of education in 1940. See Data Appendix A for details on the sample and data.

H. Empirical Bayes Shrinkage of Place Effects

When estimating the place effects with many geographic units (counties), a common problem is that some estimates may be very noisy. While these estimates are unbiased, they are on average further from the truth—in a total squared error sense—than optimal (Efron, 2010). Shrinkage techniques address this problem.

Empirical Bayes methods have become a popular means to shrink noisy estimates (e.g., Angrist et al., 2017; Chetty and Hendren, 2018). The method is motivated by the fact that under the assumption of place effects resulting from a common (unknown) distribution, the optimal point estimator has the from of a Bayesian posterior mean (Armstrong et al., 2021). One does not need to make any assumptions on the specific distribution that the place effects result from.

Figure H.14: Example of Empirical Bayes Shrinkage With a Single Covariate



Notes: This figure shows a random subsample of the 1870 ancestor county fixed effect (FE) estimates on 1940 years of education for descendants of the Enslaved. Panel (A) shows the preliminary estimates and 95 percent confidence bands clustered at the family level. Panel (B) shows the estimates after shrinking them toward the regression line based on the county's Black population share and robust 95 percent empirical Bayes confidence intervals (Armstrong et al., 2021). County FEs based on ten observations or less are discarded. See Data Appendix A for details on the sample and data.

Figure H.14 provides a simple example of empirical Bayes shrinkage with a single covariate. Panel (A) shows the negative correlation between the preliminary estimates of a county's causal effect on years of education in 1940 and county's share of the population that is Black in 1860. Panel (B) shows the empirical Bayesian shrinkage estimates. Noisier preliminary estimates are pulled toward the regression line more drastically than more precise preliminary estimates.

We employ a empirical Bayes shrinkage to our baseline county-effects. We provide two forms of shrinkage estimates. The first set does not use covariates, shrinking the baseline estimates toward a common mean. The second set includes covariates, shrinking the baseline estimates toward the place effect predicted by the covariates. We prefer the second set of estimates because they leverage a larger set of information in the shrinkage process.

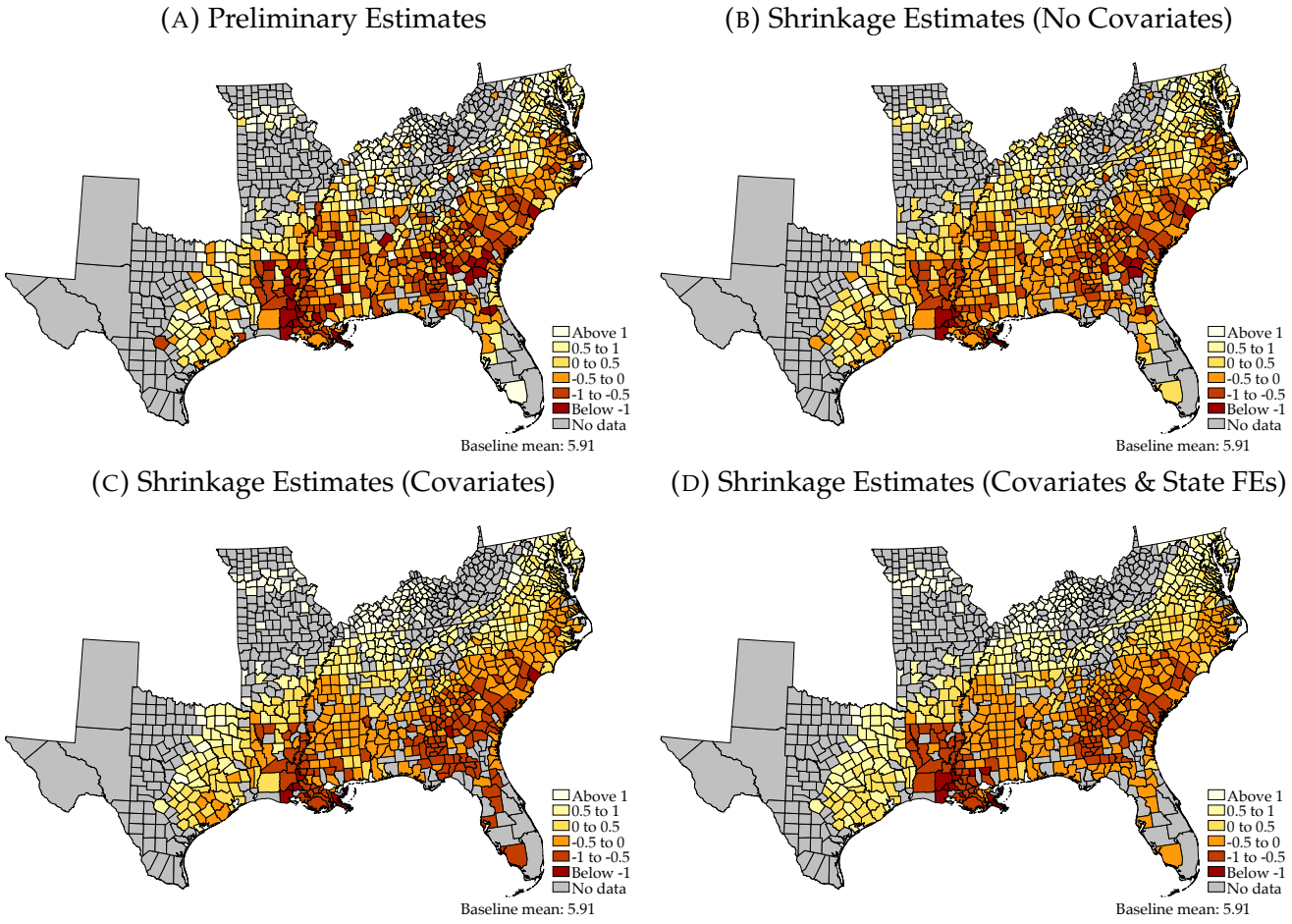
The empirical Bayes estimate for county c including covariates takes the form

$$\widehat{\theta}_c \quad = \quad X_c'\widehat{\delta} + \frac{\widehat{s}}{\widehat{s}+\widehat{\sigma}_c^2} \cdot (\eta_c - X_c'\widehat{\delta}), \tag{22}$$

where $\widehat{\eta}_c$ is the preliminary estimate of county c's effect, $\widehat{\delta} = \left(\sum_{c=1}^{N} \widehat{\sigma}_c^{-2} X_c X_c'\right)^{-1} \sum_{c=1}^{N} \widehat{\sigma}_c^{-2} X_c \widehat{\eta}_c$ is the ordinary least squares estimate of $\widehat{\eta}_c$ on the county covariates X_c, $\widehat{\sigma}_c^2$ is the standard error of $\widehat{\eta}_c$, and $\widehat{s} = \max\left\{\frac{-N+\sum_{c=1}^{N}\widehat{\sigma}_c^{-2}\widehat{\varepsilon}_c^2}{\sum_{c=1}^{N}\widehat{\sigma}_c^{-2}}, \frac{2}{\sum_{c=1}^{N}\widehat{\sigma}_c^{-2}}\right\}$ with $\widehat{\varepsilon}_c = \widehat{\eta}_c - X_c'\widehat{\delta}$. The shrinkage estimate is therefore a weighted average of the preliminary county fixed effect and the predicted effect based on the county's characteristics, with greater weight assigned to a preliminary fixed effects when it is more precisely estimated.

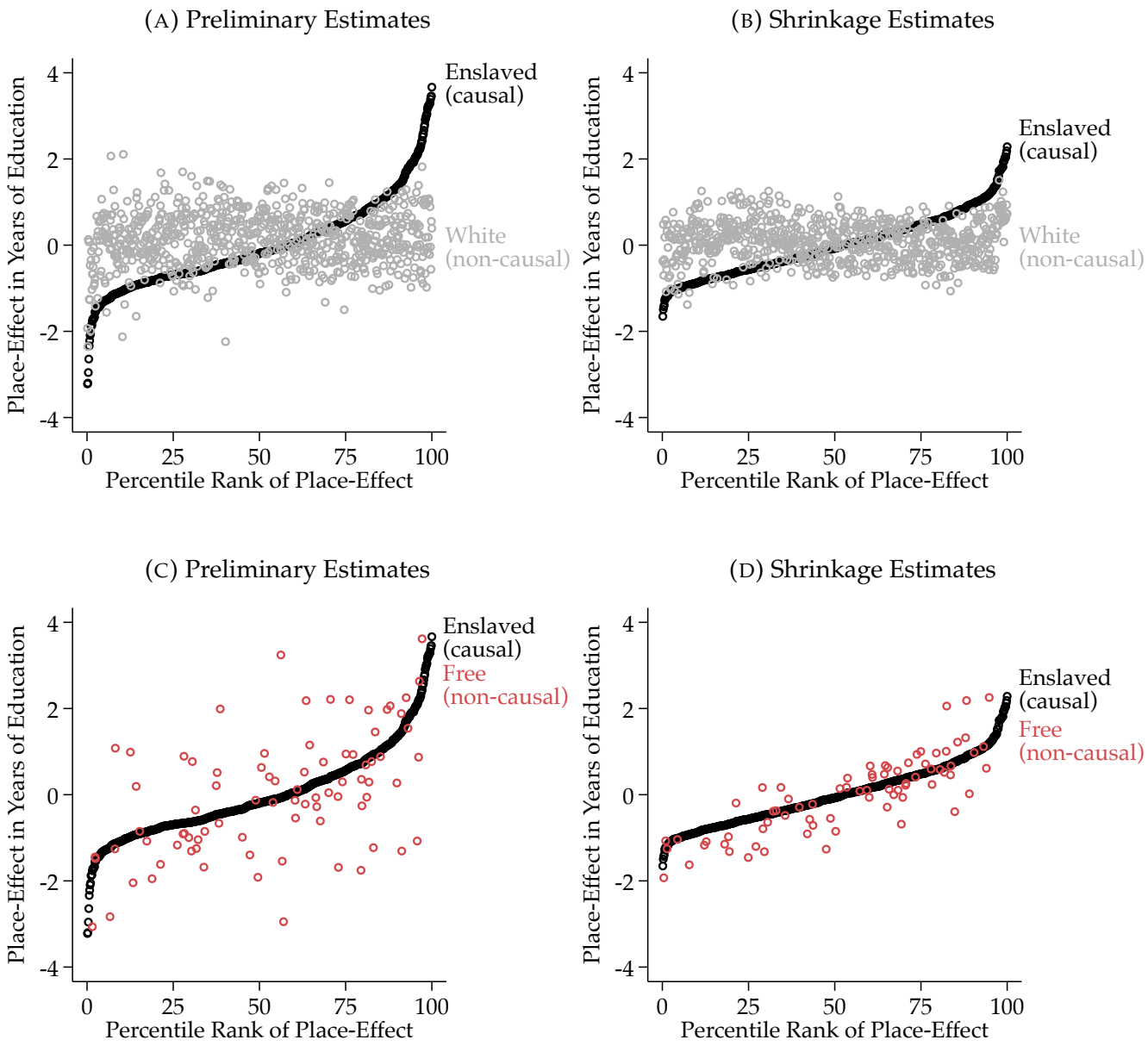
Figure H.15 shows a comparison of the place effects before and after the empirical Bayes shrinkage. The covariates we include in the shrinkage process are those listed in Figure 14. While the general pattern of negative effects being concentrated in the Deep South holds before and after, the shrunk estimates are far more spatially correlated.

FIGURE H.15: Causal Place Effects on 1940 Years of Education



Notes: This figure shows the 1870 ancestor county fixed effect (FE) estimates on 1940 years of education for descendants of the Enslaved. Panel (A) shows the preliminary estimates. Panel (B) shows the estimates after shrinking them to their common mean. Panel (C) shows the estimates after shrinking them to the regression line based on various covariates (see Figure 14). Panel (D) repeats Panel (C) adding state FEs. County FEs based on ten observations or less are discarded. See Data Appendix A for details on the sample and data.

FIGURE H.16: Place Effects Across Groups Before and After Shrinkage



Notes: This figure compares the 1870 ancestor county fixed effect estimates on years of education in 1940 for descendants of the Enslaved (causal) with those of white Americans and descendants of free Black Americans (both non-causal). Panels (A) and (C) show the estimates before shrinkage, Panels (B) and (D) show the estimates after shrinkage. The shrinkage does not preserve a county's original rank. County fixed effects based on ten observations or less are discarded. See Data Appendix A for details on the sample and data.

REFERENCES

ANGRIST, J. D., P. D. HULL, P. A. PATHAK, AND C. R. WALTERS (2017): "Leveraging Lotteries for School Value-Added: Testing and Estimation," *Quarterly Journal of Economics*, 132, 871–919.

ARMSTRONG, T. B., M. KOLESŔ, AND PLAGBORG-MØLLER (2021): "Robust Empirical Bayes Confidence Intervals," Tech. rep.

CHETTY, R. AND N. HENDREN (2018): "The impacts of neighborhoods on intergenerational mobility II: County-level estimates," *Quarterly Journal of Economics*, 133, 1163–1228.

EFRON, B. (2010): *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*, Cambridge University Press.

I. PLACEBO EXERCISES

In two types of placebo exercises, we test our method of quantifying the Free-Enslaved gap. First, we estimate the placebo Free-Enslaved gap for white Americans. White families that cannot be linked to the 1850 or 1860 censuses are classified as (placebo) descendants of the Enslaved. This exercise may not yield pure placebo estimates because families immigrating after 1860 may be different from those who immigrated earlier. Table I.4 shows that the composition of white Americans indeed experienced some changes after 1860. White Americans grew more likely to be first-generation immigrants and among those immigrants fewer came from the United Kingdom and more from Northern, Central, and Eastern Europe. When evaluating the placebo estimates, we should bare in mind that these changes contaminate a pure placebo.

Second, we estimate the Free-Enslaved gap using 1875 as the (placebo) end of slavery. Figure 1 already suggests that there is no gap between Black Americans who can be linked back to 1880 (but not 1870 or earlier) and those who can be linked back to 1870 or earlier. In this section, we also estimate the placebo Free-Enslaved gap based on the change in the distribution of last names from 1870 to 1880.

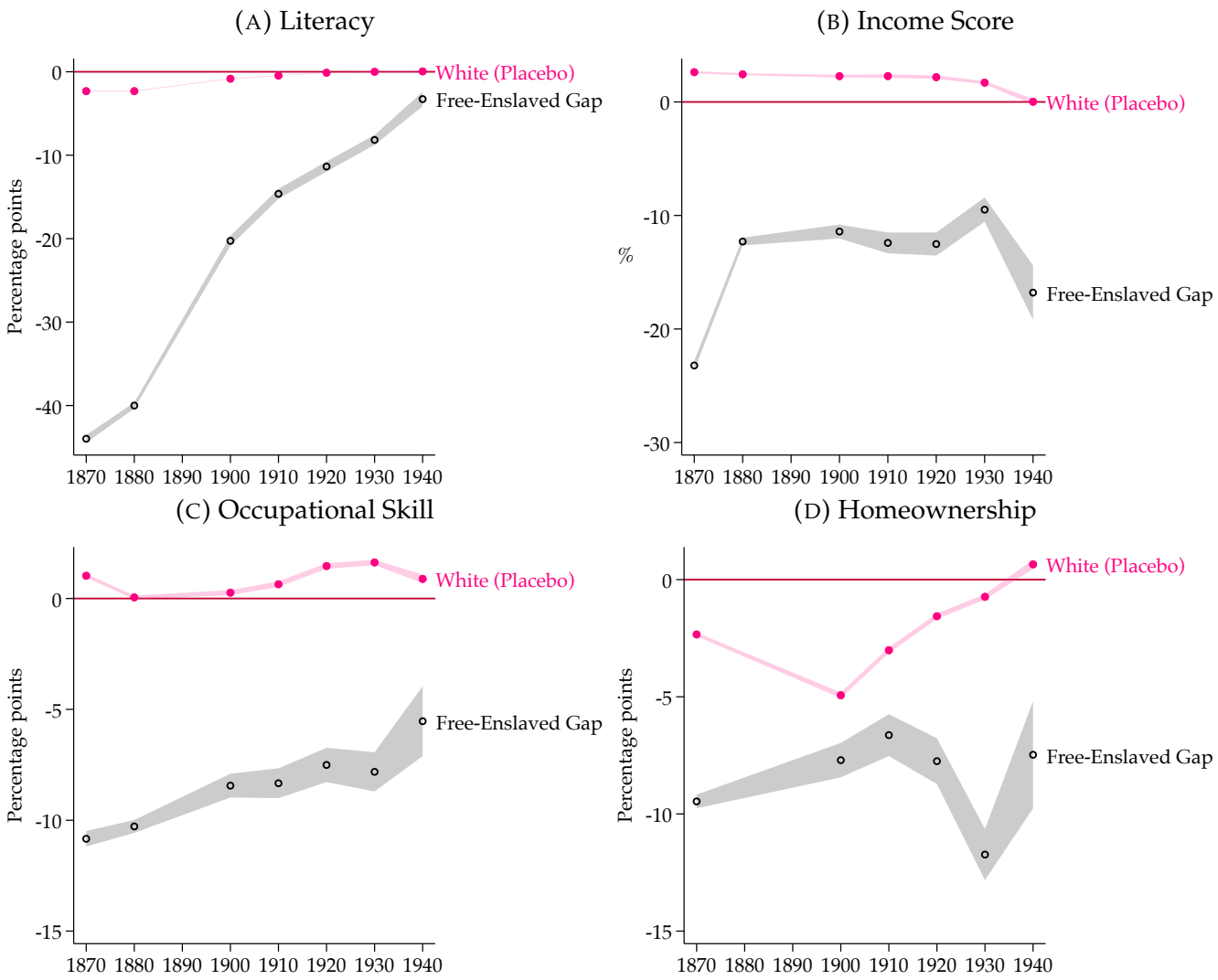
Beyond the two placebo exercises, we also estimate the effect that the 1870 ancestor birthplace has on white Americans' socioeconomic status in 1940. Because we do not have exogenous variation in the ancestor birthplace for free people such as white Americans, these estimates are non-causal. We also trace the speed of convergence in the socioeconomic status of white families whose 1870 ancestors did or did not have physical or human capital. This exercise yields a benchmark for the speed of convergence between descendants of the Enslaved and the Free. Lastly, we estimate the intergenerational effect that the Deep South's institutions had on white Americans using our border discontinuity design.

TABLE I.4: White Americans Before and After 1865

Region of Birth	Share in 1860 (%)	Share in 1870 (%)	Change (Ppts.)	Change (%)
North America	86.5	85.3	-1.2	-1.4
Central America	0.2	0.2	0.0	25.3
Northern Europe	0.2	1.0	0.8	337.3
UK	8.1	7.0	-1.0	-12.8
Western Europe	0.7	0.8	0.1	10.8
Southern Europe	0.1	0.2	0.1	76.6
Central/Eastern Europe	4.2	5.4	1.3	30.7

Notes: This table shows the share of white American men born in each region in the 1850–1860 and the 1870–1880 censuses. Regions with shares below 0.1 percent in any year are omitted.

FIGURE I.17: Free-Enslaved Gap (1870-1940) vs. Placebo for White Americans



Notes: This figure shows the true and placebo gaps in literacy rates, income scores (OCCSCORE), occupation skill levels (HISCLASS), and homeownership rates among prime-age (20-54) male descendants of free vs. enslaved Black Americans in each census decade. The placebo applies the exact same procedure to the sample of white Americans. The comparison shows that some linking bias may affect results in early periods, but all of it vanishes over time. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. All estimates control for a quadratic function in age and include 95 percent confidence bands that are clustered at the family level. While literacy is reported directly in the census decades only until 1930, we use schooling in second grade as a proxy in 1940. See Data Appendix A for details on the sample and data.

TABLE I.5: Placebo Free-Enslaved Gap (1940) for White Americans

	Education (Years) Mean: 9.76	Wage Income (USD) Mean: 892.68	Home Ownership (%) Mean: 49.74	House Value (USD) Mean: 3,284.56
Placebo	**-0.17*****	**-1.68**	**0.09**	**12.17**
	(0.00)	(1.04)	(0.05)	(9.63)
Controls (age, age^2)	Y	Y	Y	Y
Adjusted R^2	0.03	0.06	0.01	0.00
Observations	5,015,270	4,770,969	5,012,884	2,425,204
Ancestor Free	*3,158,604*	*3,001,138*	*3,155,980*	*1,536,909*

Notes: This table shows the placebo gaps in years of education, total income, homeownership, and house value among prime-age (20-54) male Black Americans in 1940. The placebo applies the exact same procedure to the sample of white Americans. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE I.6: Placebo vs. Free-Enslaved Gap Based on the Distribution of Last Names (1940)

	Education (Years)				
P(Ancestor Enslaved)			**-0.47*****	**-0.43*****	**-0.54*****
			(0.02)	(0.02)	(0.02)
Placebo: 1875	-0.05***		-0.03***		0.21***
	(0.01)		(0.01)		(0.01)
Placebo: White		-0.04***		-0.04***	0.07***
		(0.00)		(0.00)	(0.00)
Placebo: 1875×White					-0.24***
					(0.00)
Name-measure	Exact	Exact	Exact	Exact	Exact
Controls (race, race×age, race×age^2)	Y	Y	Y	Y	Y
Adjusted R^2	0.03	0.03	0.03	0.14	0.14
Observations	2,859,747	24,267,079	2,859,747	27,126,826	27,126,826

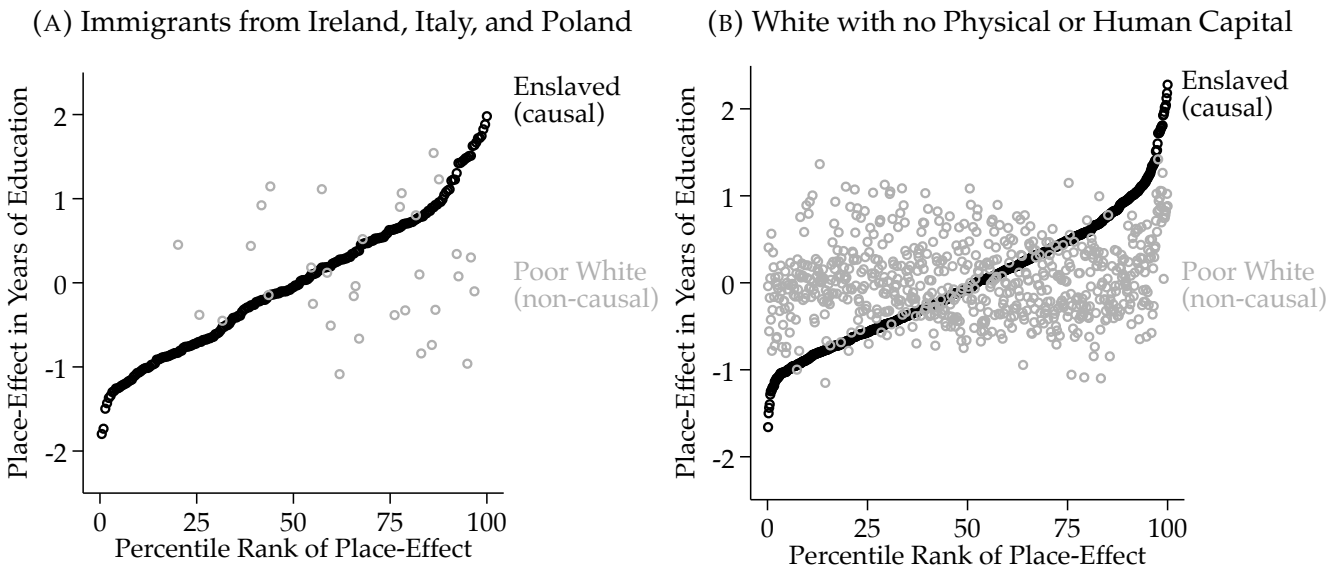
Notes: This table repeats Table 1 showing the gap in years of education, total income, homeownership, and house value among prime-age (20-54) male descendants of free vs. enslaved Black Americans in 1940. The sample includes the entire universe of prime-age Black men, not just those linkable. The coefficients can be interpreted as a 100 percentage point increase in the likelihood of descending from the Enslaved based on their (exact) last name. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Robust standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE I.7: Placebo Free-Enslaved Gap (2000) at Census Tract Level

	HS Degree (%) Mean: 83.52	College Degree (%) Mean: 25.52	Income (USD) Mean: 46,123.73	House Value (USD) Mean: 133,616.06
Ancestor Enslaved	**-0.11****	**0.22*****	**133.37**	**908.63***
	(0.05)	(0.08)	(91.00)	(499.87)
Level	Tract×Race×Sex	Tract×Race×Sex	Tract×Race	Tract×Race
Controls (age, age^2)	Y	Y	Y	Y
Adjusted R^2	0.00	0.01	0.00	0.00
Observations	437,099	437,099	437,076	435,809
Ancestor Free	*383,221*	*383,221*	*383,213*	*382,049*

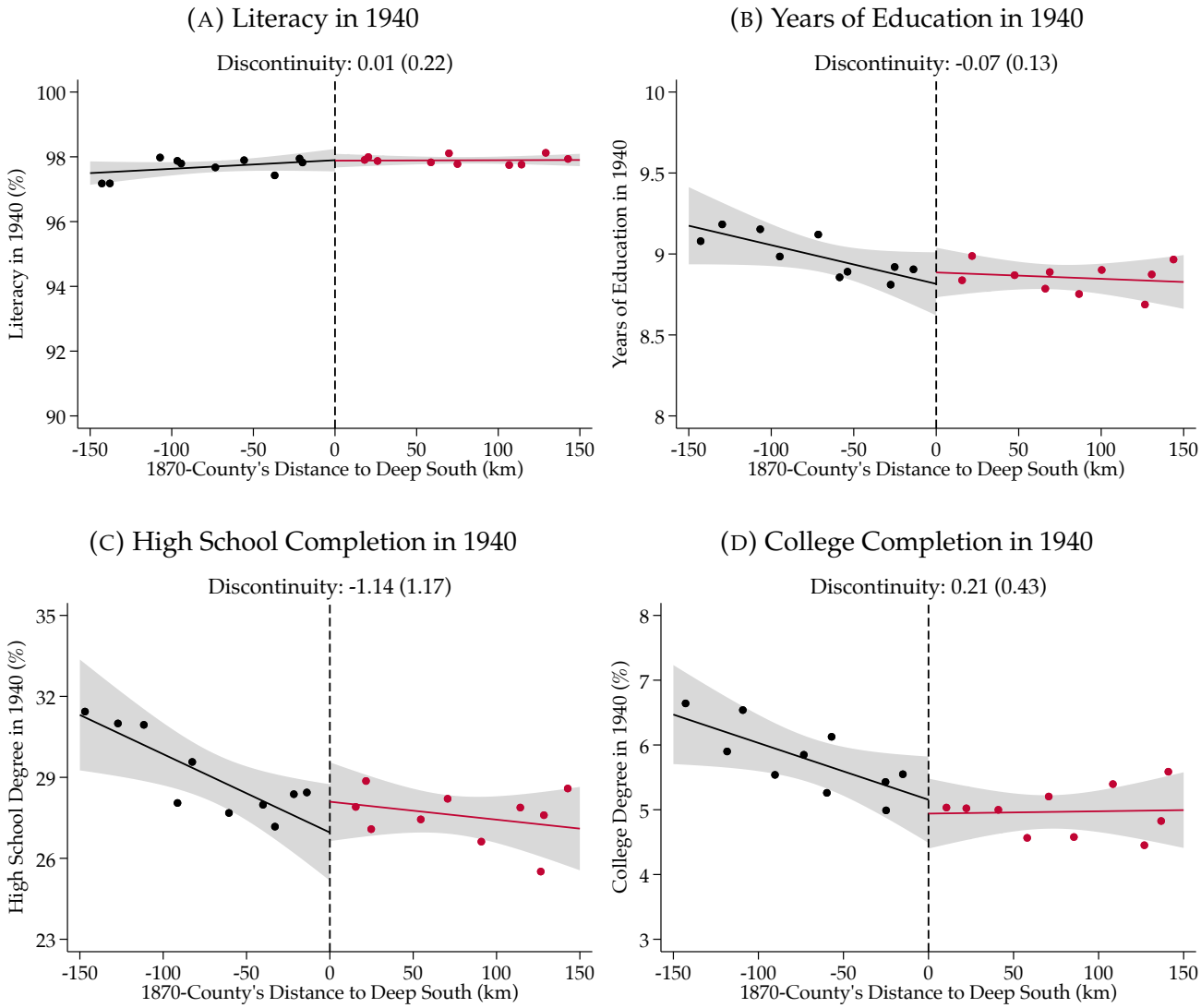
Notes: This table shows the placebo gap in neighborhood-level outcomes among male white Americans around 2000. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. We estimate the gap in terms of the fraction of people who hold a high school or college degree, the median income earned, and the median house value. House values are measured conditional on ownership. Each person is assigned the respective value of the census tract in which they lived at the time of death. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

FIGURE I.18: Place Effects Across Poor Groups



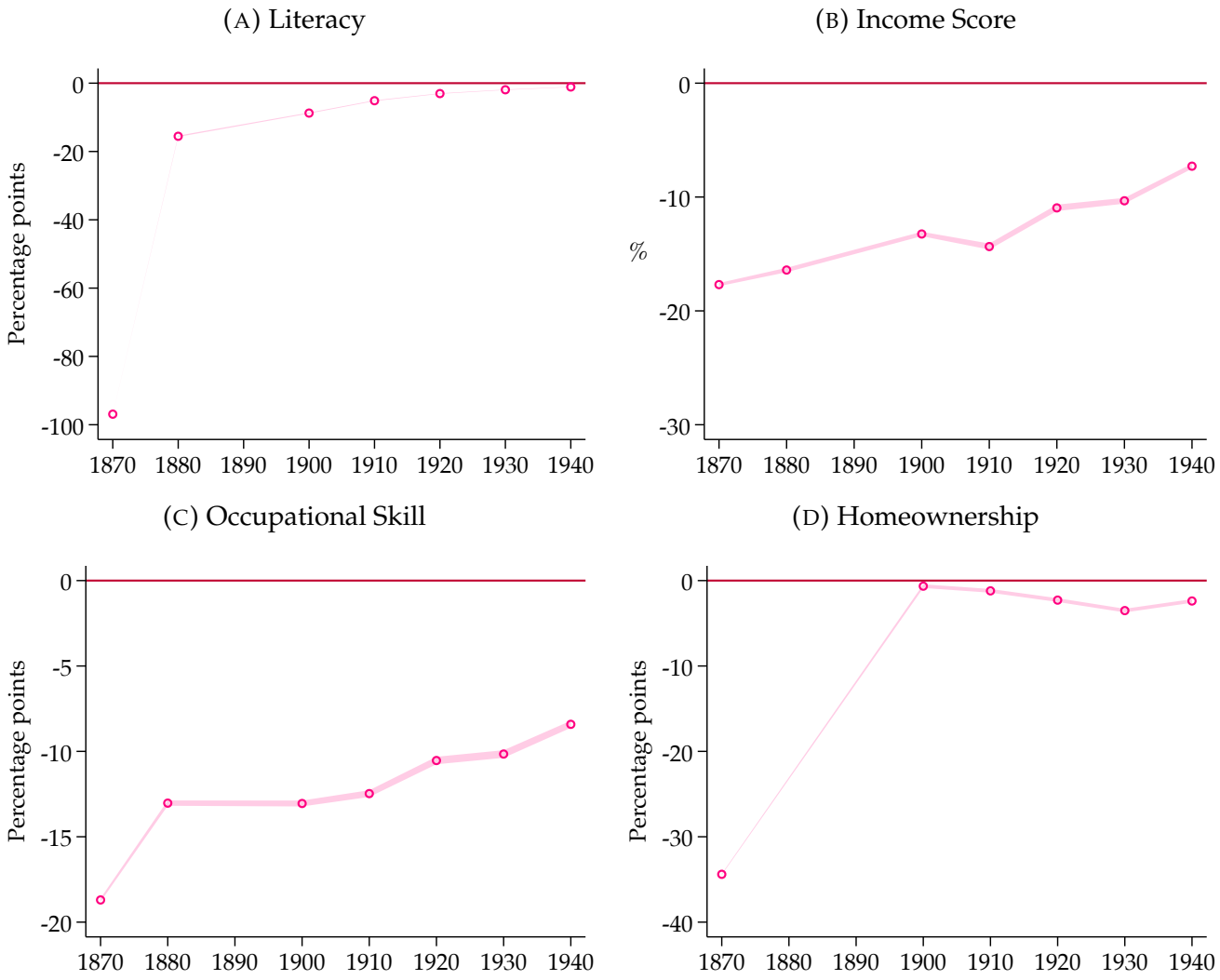
Notes: This figure compares the 1870 ancestor county fixed effect estimates on years of education in 1940 for descendants of the Enslaved (causal) with those of white Americans who descend from Irish, Italian, or Polish immigrants ($\rho = -0.041$, Panel A) and white Americans whose ancestors were illiterate and had zero wealth in 1870 ($\rho = 0.151$, Panel B). A county's fixed effect is the deviation from the race-specific population-weighted average across all Southern counties after controlling for a quadratic function of age. The sample includes only prime-age (20–54) men whose ancestors can be located in 1870. All county fixed effects based on ten observations or less are discarded. See Data Appendix A for details on the sample and data.

FIGURE I.19: No Border Discontinuities among White Americans



Notes: This figure shows a binned scatter plots of years of education, literacy, high school completion, and college completion among white prime-age (20-54) men in 1940 by their ancestor's distance to the border of the Deep South in 1870. Lines represent the best linear fit of the individual-level data. Shaded areas represent 95 percent level confidence bands with clustering at the county level. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from the county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

FIGURE I.20: Benchmark for Speed of Convergence—White Americans Whose Ancestors Did vs. Did Not Have Any Physical or Human Capital



Notes: This figure shows the gaps in literacy, income, skill, and homeownership among white prime-age (20-54) male descendants of ancestors with vs. without any physical or human capital in 1870. Physical capital is measured in terms of real and personal property; Human capital is measured in terms of literacy. The comparison yields a benchmark for the convergence of large socioeconomic gaps from 1870 to 1940. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. All estimates control for a quadratic function in age and include 95 percent confidence bands that are clustered at the family level. While literacy is reported directly in the census decades only until 1930, we use schooling in second grade as a proxy in 1940. See Data Appendix A for details on the sample and data.

J. Additional Results

J.1 The Free-Enslaved Gap in the 21st Century

TABLE J.8: Free-Enslaved Gap (2000) at Census Tract Level

	HS Degree (%) Mean: 68.85	College Degree (%) Mean: 12.31	Income (USD) Mean: 29,875.58	House Value (USD) Mean: 87,921.78
Ancestor Enslaved	**-3.02*****	**-2.45*****	**-4,795.93*****	**-15,755.30*****
	(0.51)	(0.55)	(636.79)	(2,462.82)
Level	Tract×Race×Sex	Tract×Race×Sex	Tract×Race	Tract×Race
Controls (age, age^2)	Y	Y	Y	Y
Adjusted R^2	0.01	0.00	0.01	0.00
Observations	26,765	26,765	26,803	25,787
Ancestor Free	*1,713*	*1,713*	*1,715*	*1,634*

Notes: This table shows the gap in neighborhood-level outcomes among descendants of free vs. enslaved Black Americans around 2000. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. We estimate the gap in terms of the fraction of people with a high school or college degree, their median income, and their median house value. House values are measured conditional on ownership. Each person is assigned the respective value of the census tract in which they lived at the time of death. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE J.9: Free-Enslaved Gap (2000) at Five-Digit ZIP Code Level

	HS Degree (%) Mean: 68.50	College Degree (%) Mean: 11.86	Income (USD) Mean: 28,227.95	House Value (USD) Mean: 95,696.55
Ancestor Enslaved	**-3.51*****	**-2.34*****	**-4,563.02*****	**-14,730.17*****
	(0.40)	(0.32)	(461.43)	(1,795.31)
Level	ZIP×Race×Sex	ZIP×Race×Sex	ZIP×Race	ZIP
Controls (age, age^2)	Y	Y	Y	Y
Adjusted R^2	0.01	0.00	0.01	0.00
Observations	31,465	31,465	31,480	31,760
Ancestor Free	*1,972*	*1,972*	*1,973*	*1,992*

Notes: This table shows the gap in neighborhood-level outcomes among descendants of free vs. enslaved Black Americans around 2000. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. We estimate the gap in terms of the fraction of people with a high school or college degree, their median income, and their median house value. House values are measured conditional on ownership. Each person is assigned the respective value of the five-digit ZIP code in which they lived at the time of death. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

J.2 The Free-Enslaved Gap in Alternative Measures

TABLE J.10: Free-Enslaved Gap (1940) in Different Income Measures

	OCCSCORE (1950-USD) Mean: 1,604.09	LIDO Score (1950-USD) Mean: 1,161.69	Wage Income (1940-USD) Mean: 381.20	Total Income (1940-USD) Mean: 793.47
Ancestor Enslaved	**-148.39*****	**-279.00*****	**-145.92*****	**-204.29*****
	(10.86)	(8.59)	(6.13)	(10.29)
Controls (age, age^2)	Y	Y	Y	Y
Adjusted R^2	0.04	0.04	0.05	0.09
Observations	168,138	142,743	154,463	146,871
Ancestor Free	*9,325*	*7,517*	*8,551*	*8,100*

Notes: This table shows the Free-Enslaved gap in income across different measures: Occupational income score (OCCSCORE), a refined occupational income score (Lido), wage income, and total predicted income. All estimates are for Black prime-age men in 1940. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE J.11: Free-Enslaved Gap (1940) in Different Income Measures

	Literacy (%) Mean: 91.49	Education (Years) Mean: 5.99	High School (%) Mean: 9.28	College (%) Mean: 1.70	Graduate (%) Mean: 0.46
Ancestor Enslaved	**-4.25*****	**-1.59*****	**-7.86*****	**-1.86*****	**-0.74*****
	(0.26)	(0.05)	(0.45)	(0.21)	(0.12)
Controls (age, age^2)	Y	Y	Y	Y	Y
Adjusted R^2	0.01	0.04	0.01	0.00	0.00
Observations	163,549	163,549	163,549	163,549	163,549
Ancestor Free	*9,078*	*9,078*	*9,078*	*9,078*	*9,078*

Notes: This table shows the Free-Enslaved gap in education across different measures: Literacy, years of education, and the probability of holding a high school, college, or graduate degree. As literacy is reported only until 1930, we use educational attainment beyond second grade as a proxy in 1940. All estimates are for Black prime-age men in 1940. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

J.3 The Free-Enslaved Gap in Intergenerational Mobility

To determine if the persisting Free-Enslaved gap is purely the result of large initial differences or if are there also differences in intergenerational mobility, we estimate the intergenerational mobility of person i between 1870 and 1940 as the following:

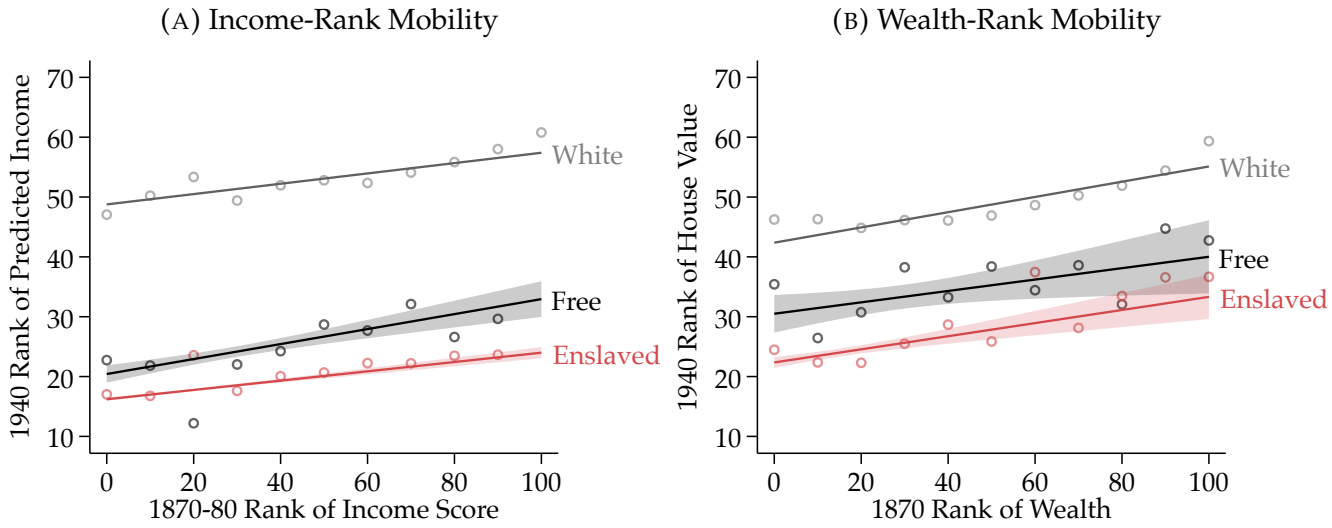
$$r_{i,1940} = \alpha + \beta_1 s_i + \beta_2 r_{i,1870} + \beta_3 s_i \times r_{i,1870} + \varepsilon_i, \tag{23}$$

where $r_{i,1940}$ is the percentile rank of i's 1940 income or wealth in the Black distribution and $r_{i,1870}$ is the percentile rank of the income or wealth occupied by the prime-age male ancestor of i in

1870. Upward mobility is defined as the increase in percentile ranks from one generation to the next. To reduce the impact of measurement error, we use the average across 1870 and 1880 for ancestors if possible (**?**).

Figure J.21 shows that upward mobility is significantly lower for descendants of the Enslaved. Across all parts of the 1870–1880 income distribution and the 1870 wealth distribution, descendants of the Enslaved reach lower expected income and wealth levels in 1940.

FIGURE J.21: Gaps in Intergenerational Mobility (1870-1940)



Notes: This figure shows the estimated intergenerational mobility of free Black (in black), formerly enslaved (in red), and white prime-age (20-54) men. The left panel shows income-rank mobility conditional on the average rank of an ancestor's occupational income score (OCCSCORE) in 1870 and 1880 (**?**). The right panel shows wealth-rank mobility conditional on the rank of an ancestor's wealth (personal property and real property) in 1870. 1940 total income is predicted (see Appendix A.5). 1940 house values are measured conditional on ownership. The lines reflect the coefficients estimated in Equation 23, including robust confidence bands at the 95 percent level. Bubbles reflect binned averages. Percentile ranks reflect rank that each income occupies in the year-specific national income distribution of Black prime-age men. See Data Appendix A for details on the sample and data.

J.4 The Free-Enslaved Gap Based on the Distribution of Last Names

TABLE J.12: Free-Enslaved Gap Based on the Distribution of Last Names (1940)

	Education (Years) Mean: 5.71		Wage Income (USD) Mean: 598.74		Homeownership (%) Mean: 21.89		House Value (USD) Mean: 1,599.75	
P(Ancestor Enslaved)	**-0.47*****	**-0.54*****	**-13.73*****	**-29.89*****	**-2.43*****	**-2.43*****	**-630.53****	**-970.17***
	(0.02)	(0.02)	(5.26)	(7.17)	(0.21)	(0.29)	(277.72)	(506.45)
Name-measure	Exact	NYSIIS	Exact	NYSIIS	Exact	NYSIIS	Exact	NYSIIS
Controls (age, age^2)	Y	Y	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.03	0.03	0.00	0.00	0.01	0.01	0.00	0.00
Observations	2,859,747		2,821,235		2,842,572		601,789	

Notes: This table repeats Table 1 showing the gap in years of education, total income, homeownership, and house value among prime-age (20-54) male descendants of free vs. enslaved Black Americans in 1940. The sample includes the entire universe of prime-age Black men, not just those linkable. The coefficients can be interpreted as a 100 percentage point increase in the likelihood of descending from the Enslaved based on their (exact) last name. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Robust standard errors are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

J.5 The Free-Enslaved Gap for Free Without Physical and Human Capital

TABLE J.13: Free-Enslaved Gap (1940) for Free Without Physical and Human Capital in 1860

	Education (Years) Mean: 5.83		Wage Income (USD) Mean: 381.64		Homeownership (%) Mean: 29.08		House Value (USD) Mean: 1,380.43	
Ancestor Enslaved	**-1.00*****	**-0.12**	**-90.43*****	**26.85**	**-6.16*****	**-1.42**	**-343.74****	**440.28****
	(0.15)	(0.15)	(21.13)	(21.44)	(1.95)	(2.00)	(159.58)	(184.15)
1870 State of Birth-FE	N	Y	N	Y	N	Y	N	Y
Controls (age, age^2)	Y	Y	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.03	0.07	0.04	0.07	0.01	0.02	0.00	0.03
Observations	71,574	71,574	67,672	67,672	72,013	72,013	20,455	20,455
Ancestor Free	*608*	*608*	*569*	*569*	*605*	*605*	*206*	*206*

Notes: This table shows the gap in years of education, total income, homeownership rate, and house value among prime-age (20-54) male descendants of a subset of the free vs. enslaved Black Americans in 1940. Among the Free, we only include those whose ancestors had no measurable physical capital (real and personal property) or human capital (literacy) in 1860. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. Columns 1, 3, 5, and 7 repeat Table 1 but hold the sample constant to the other columns. Columns 2, 4, 6, and 8 add fixed effects for 1870 ancestor state of birth. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and the Enslaved. Appendix Figure 5 shows the evolution of the conditional Free-Enslaved gap over time. See Data Appendix A for details. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

J.6 The Free-Enslaved Gap between and within Ancestor's Birthplace

TABLE J.14: Free-Enslaved Gap (1940) between and within Ancestor's Birthplace

	Education (Years) Mean: 5.91		Wage Income (USD) Mean: 388.01		Home Ownership (%) Mean: 29.48		House Value (USD) Mean: 1,412.17	
Ancestor Enslaved	-1.49***	**-0.41*****	-137.00***	**-20.22****	-6.76***	**-1.61**	-574.06***	**8.40**
	(0.07)	(0.08)	(8.51)	(9.84)	(0.86)	(1.04)	(90.08)	(115.61)
1870 State of Birth-FE	N	Y	N	Y	N	Y	N	Y
Controls (age, age^2)	Y	Y	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.04	0.08	0.04	0.07	0.01	0.03	0.01	0.03
Observations	75,583	75,583	71,474	71,474	76,048	76,048	21,873	21,873
Ancestor Free	*4,617*	*4,617*	*4,371*	*4,371*	*4,640*	*4,640*	*1,624*	*1,624*

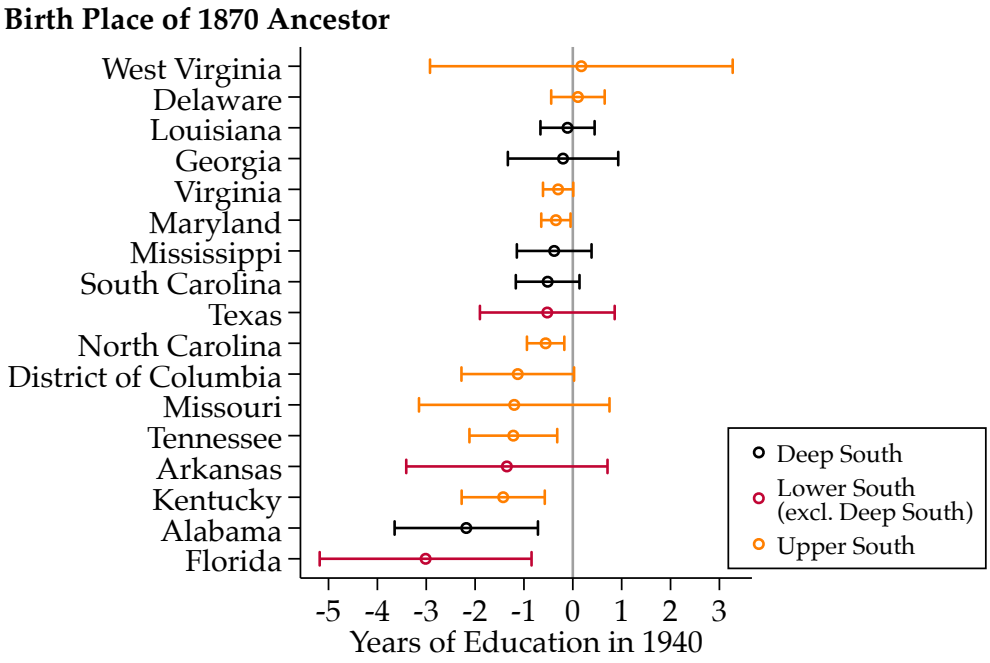
Notes: This table shows the gap in years of education, total income, homeownership rate, and house value among prime-age (20-54) male descendants of free vs. enslaved Black Americans in 1940. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. Columns 1, 3, 5, and 7 repeat Table 1 but hold the sample constant to the other columns. Columns 2, 4, 6, and 8 add fixed effects for 1870 ancestor state of birth. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and the Enslaved. Appendix Figure 5 shows the evolution of the conditional Free-Enslaved gap over time. See Data Appendix A for details. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE J.15: Free-Enslaved Gap (2000) at Census Block Level between and within Ancestor's Birthplace

	HS Degree (%) Mean: 69.20		College Degree (%) Mean: 12.32		Income (USD) Mean: 30,143.90		House Value (USD) Mean: 88,830.12	
Ancestor Enslaved	-2.57***	**-0.89**	-2.07***	**-0.29**	-5,032.50***	**-1,014.92**	-13,391.02***	**-780.04**
	(0.74)	(0.82)	(0.78)	(0.78)	(921.89)	(1,005.32)	(3,498.95)	(3,829.19)
Level	Tract×Race×Sex		Tract×Race×Sex		Tract×Race		Tract×Race	
1870 State of Birth-FE	N	Y	N	Y	N	Y	N	Y
Controls (age, age^2)	Y	Y	Y	Y	Y	Y	Y	Y
Adjusted R^2	0.00	0.01	0.00	0.01	0.01	0.03	0.00	0.02
Observations	11,931	11,931	11,931	11,931	11,932	11,932	11,500	11,500
Ancestor Free	*863*	*863*	*863*	*863*	*861*	*861*	*830*	*830*

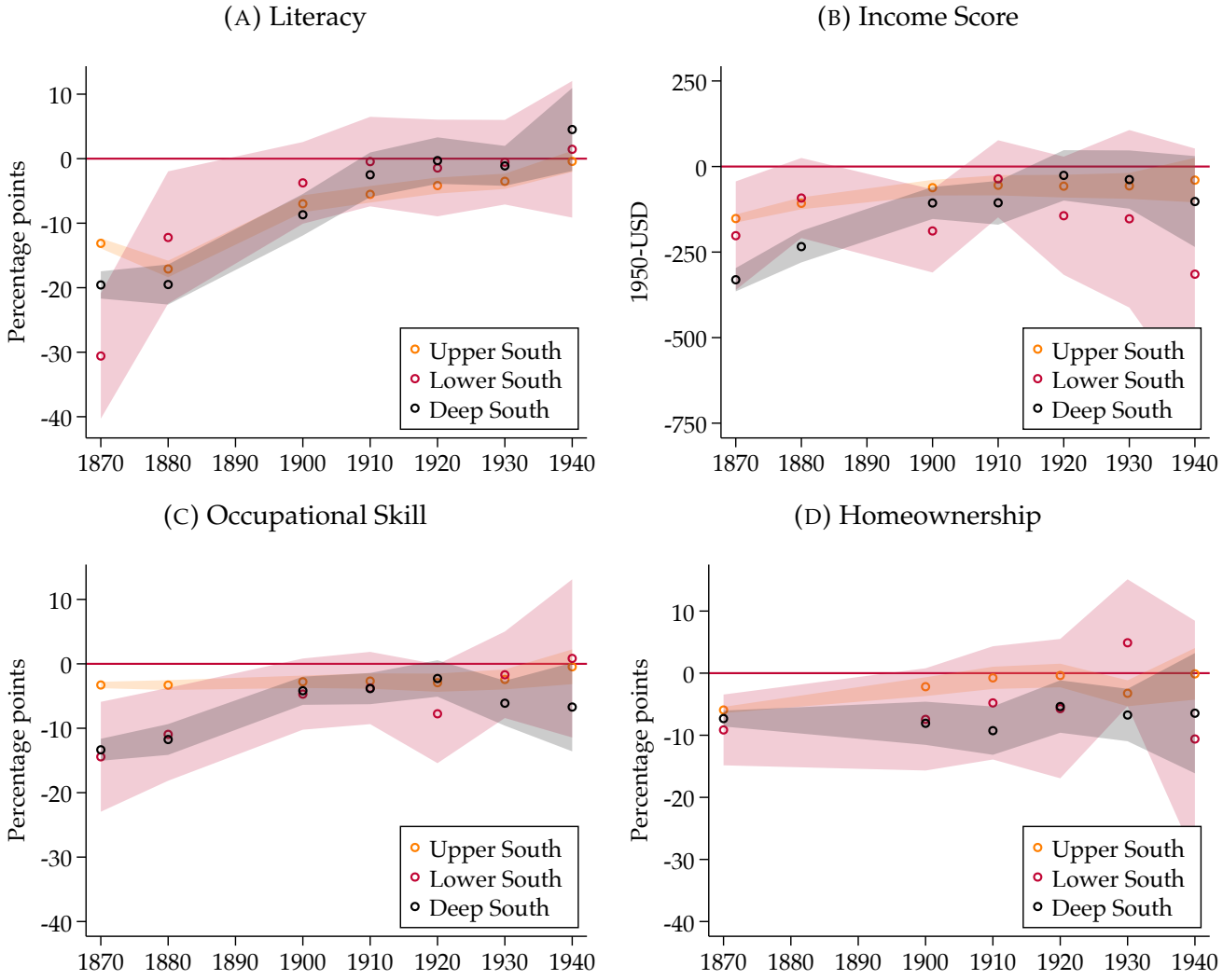
Notes: This table shows the Free-Enslaved gap in the fraction of people who hold a high school degree, the fraction of people who hold a college degree, the median income earned, and the median house value in 2000. Columns 1, 3, 5, and 7 repeat Table J.9 but hold the sample constant to the other columns. Columns 2, 4, 6, and 8 add fixed effects for 1870 ancestor state of birth. House values are measured conditional on ownership and therefore exclude zeros. Each person is assigned the respective value of the census block in which they lived at the time of death. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

FIGURE J.22: Free-Enslaved Gap in 1940 Years of Education by 1870 Ancestor Birthplace



Notes: This figure shows the gaps between descendants of Free and Enslaved in 1940 years of education by 1870 ancestor state of birth. The comparison is made between prime-age (20-54 years) male descendants in each census decade. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included, minimizing bias due to the fact that the Free by definition have a link to 1850 or 1860. Both panels control for age and include 95 percent confidence bands that are clustered at the family level.

FIGURE J.23: Free-Enslaved Gap (1870-1940) by Ancestor Region



Notes: This figure shows the gaps in literacy, income, skill, and homeownership rates among prime-age (20-54) male descendants of free vs. enslaved Black Americans in each census decade by region. The region is determined by the state of birth of a person's 1870 ancestor. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. All estimates control for a quadratic function in age and include 95 percent confidence bands that are clustered at the family level. While literacy is reported directly in the census decades only until 1930, we use schooling in second grade as a proxy in 1940. See Data Appendix A for details on the sample and data.

J.7 The Free-Enslaved Gap in Mortality

TABLE J.16: Mortality Gaps between Descendants of the Free and the Enslaved (2000)

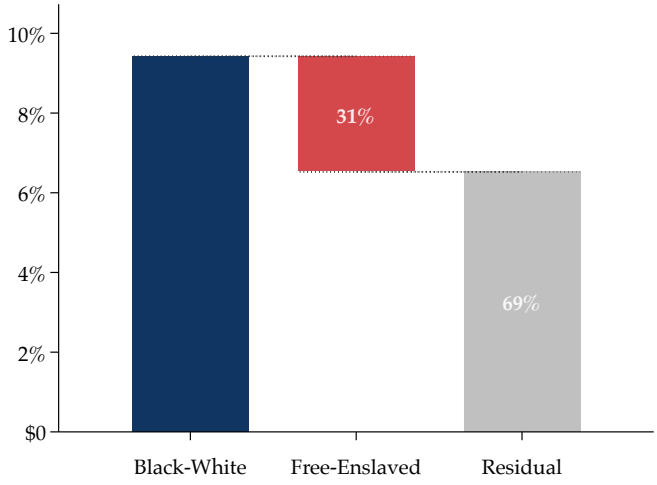
	Life Expectancy (Years)	Fatal Police Encounters (per 100K)
	Mean: 74.94	Mean: 21.31
Ancestor Enslaved	**-1.08*****	**3.40*****
	(0.14)	(1.19)
Level	Tract	ZIP
Controls (age, age^2)	Y	Y
Adjusted R^2	0.00	0.00
Observations	26,181	25,246
Ancestor Free	*1,686*	*1,423*

Notes: This table shows the Free-Enslaved gap in life expectancy and fatal police encounters per 100,000 residents. Each person is assigned the respective value of the census tract or five-digit ZIP code in which they lived at the time of death. Sample means are computed for the combined sample of the Free and the Enslaved. See Data Appendix A for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

J.8 Decomposing the Black-White Gap in Education

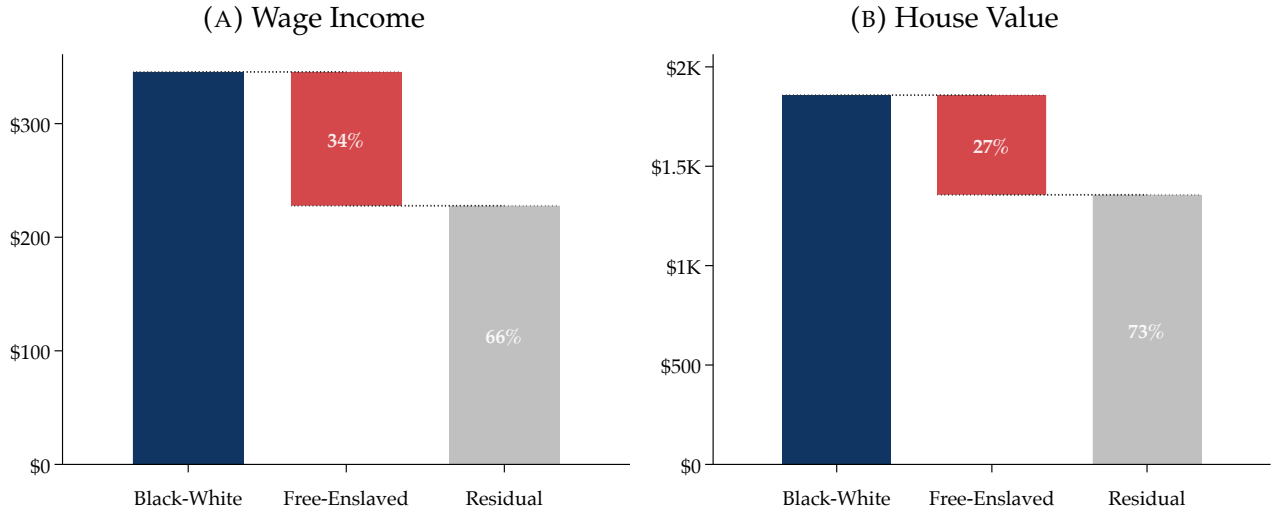
FIGURE J.24: Decomposition of the Black-White College Gap (2000) at Census Tract Level

(A) College Degree



Notes: This figure decomposes the Black-white gap in college attainment into the component accounted for by the Free-Enslaved gap and a residual. For this decomposition, we only use variation across census tract. Each person is assigned the respective race- and gender-specific median across the census tract in which they lived at the time of death. See Data Appendix A for details on the sample and data.

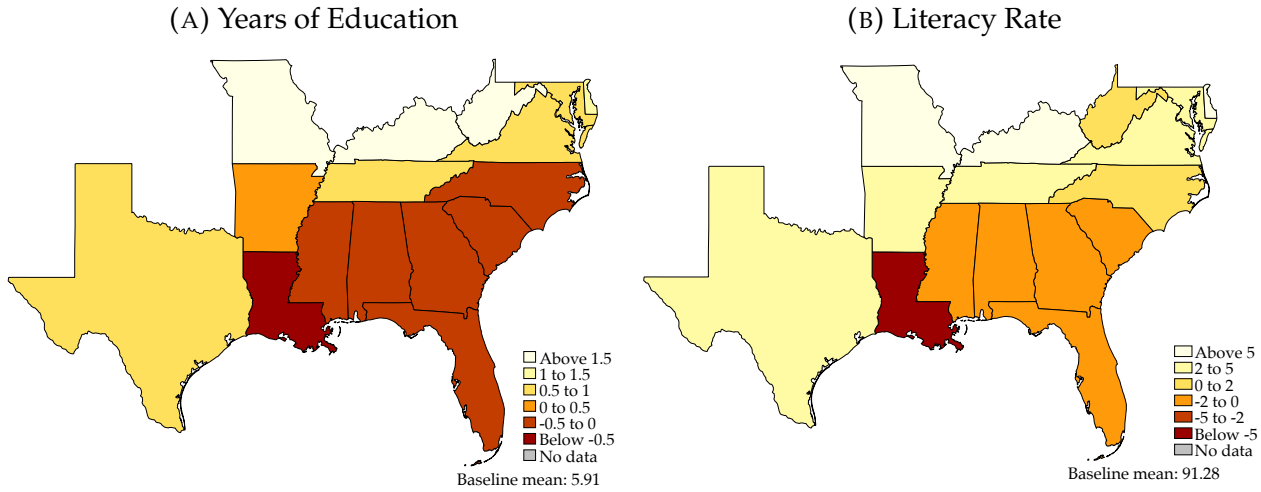
FIGURE J.25: Decomposition of the Black-White Gap (1940) at Individual Level



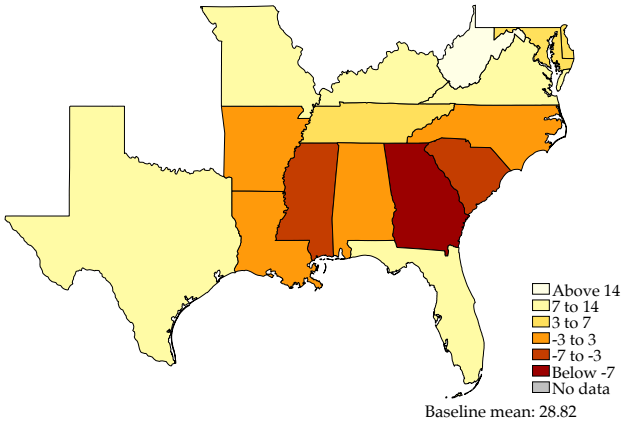
Notes: This figure decomposes the Black-white gap among prime-age (20–54) men in 1940 in wage income (Panel A) and house values (Panel B) into the component accounted for by the Free-Enslaved gap and a residual. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. House values are measured conditional on ownership. See Data Appendix A for details on the sample and data.

J.9 The Intergenerational Effect of Exposure to State-Specific Factors

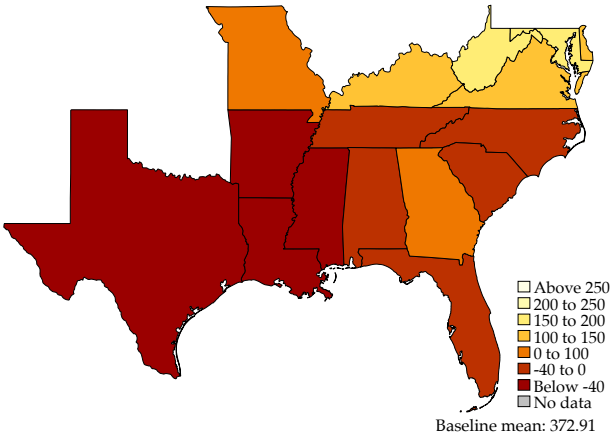
FIGURE J.26: Effect of 1870 Ancestor's State of Birth on 1940 Outcomes



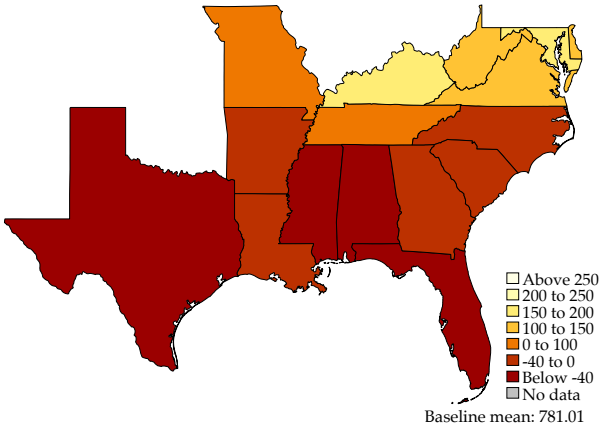
(C) Homeownership Rate



(D) Wage Income



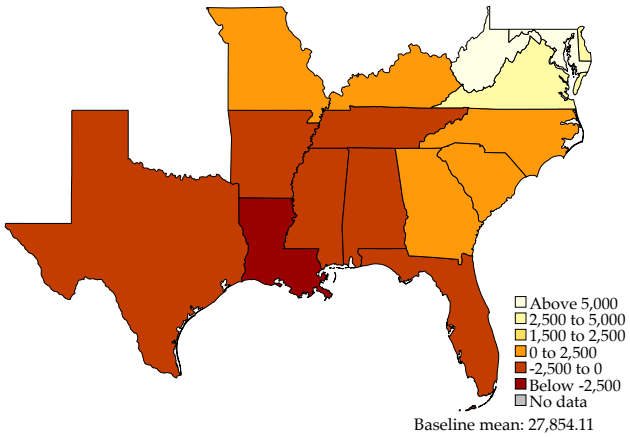
(E) Total Predicted Income



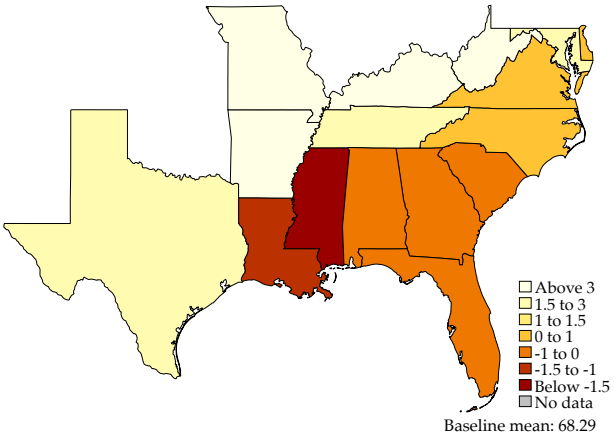
Notes: This figure shows the 1870 ancestor state of birth fixed effect estimates on years of education, literacy rate, homeownership rate, wage income, and predicted total income in 1940. A state's fixed effect is the deviation from the population-weighted average across all states after controlling for a quadratic function of age. The sample includes only Black prime-age men whose ancestors can be located in 1870.

FIGURE J.27: Effect of Ancestor's State of Birth of Enslavement on 2000 Outcomes

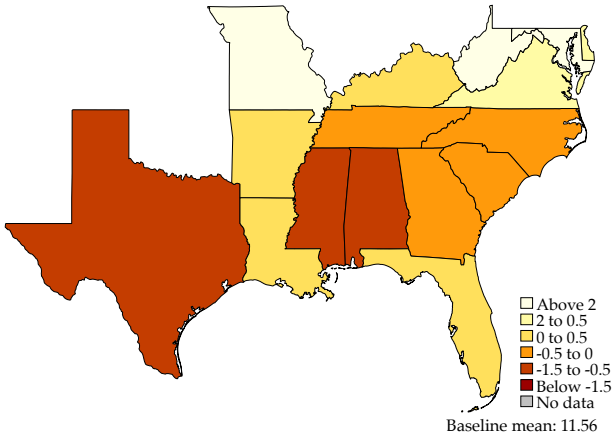
(A) Income (2000 USD)



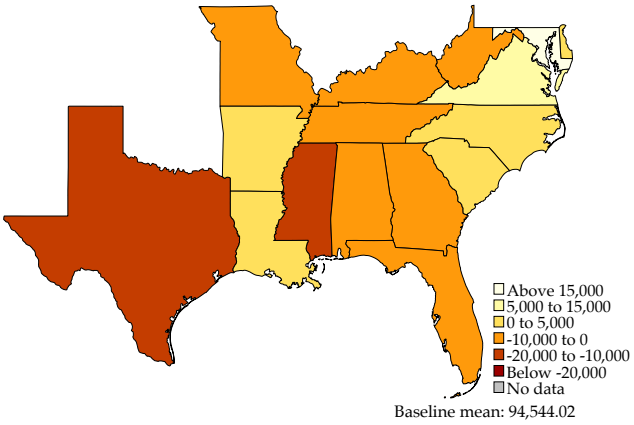
(B) High School Degree (%)



(C) College Degree (%)



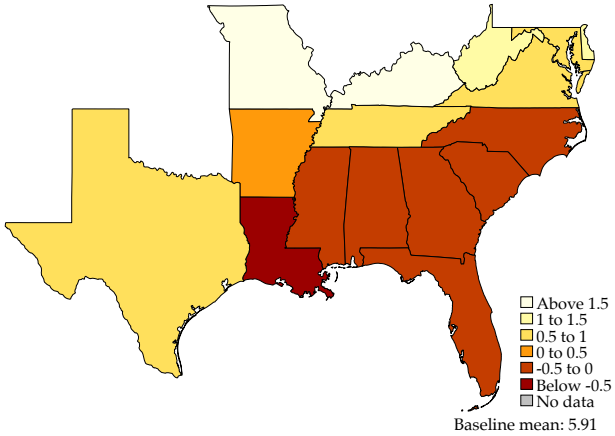
(D) House Value (2000 USD)



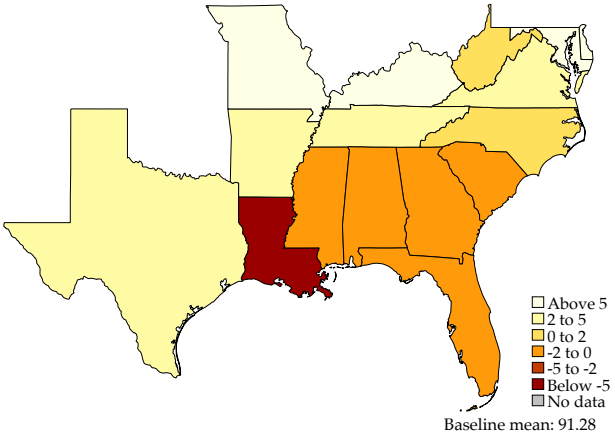
Notes: This figure shows the 1870 ancestor state of birth fixed effect estimates on income, the likelihood of holding a high school degree or college degree, and house values in 2000 for descendants of the Enslaved. A state's fixed effect is the deviation from the population-weighted average across all states (baseline mean) after controlling for a quadratic function of age. The sample includes men and women whose ancestors can be located in 1870. All county fixed effects based on ten observations or less are discarded.

FIGURE J.28: Effect of 1870 Ancestor's State of *Residence* on 1940 Outcomes

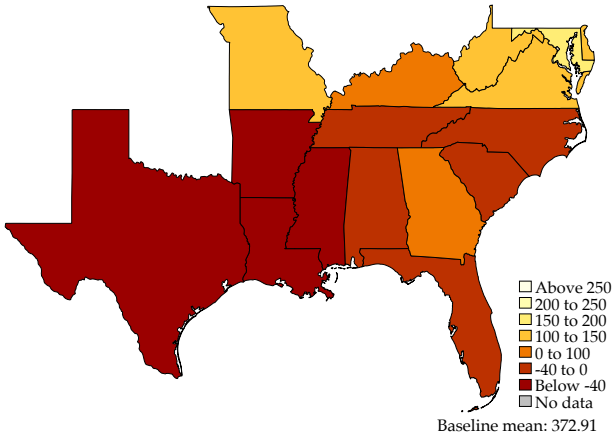
(A) Years of Education



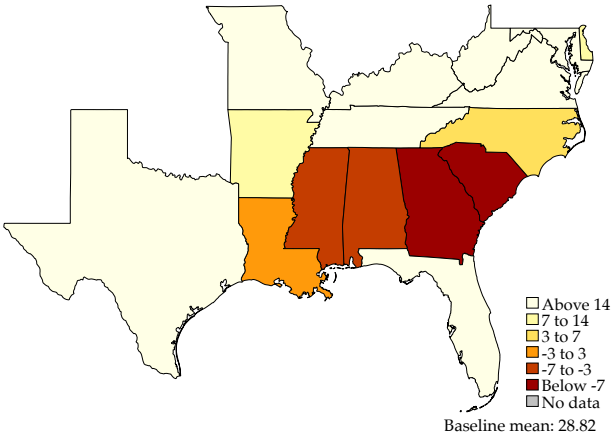
(B) Literacy Rate



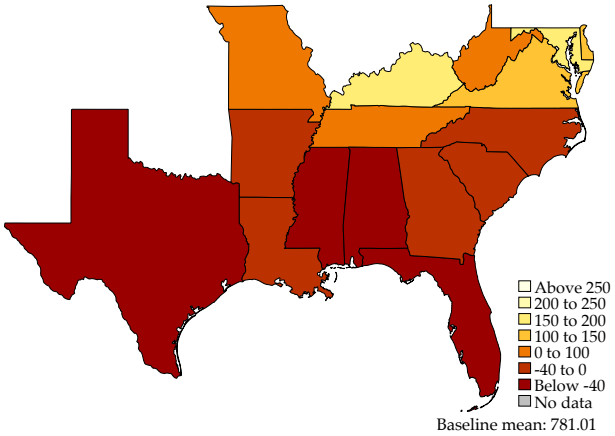
(C) Wage Income



(D) Homeownership Rate



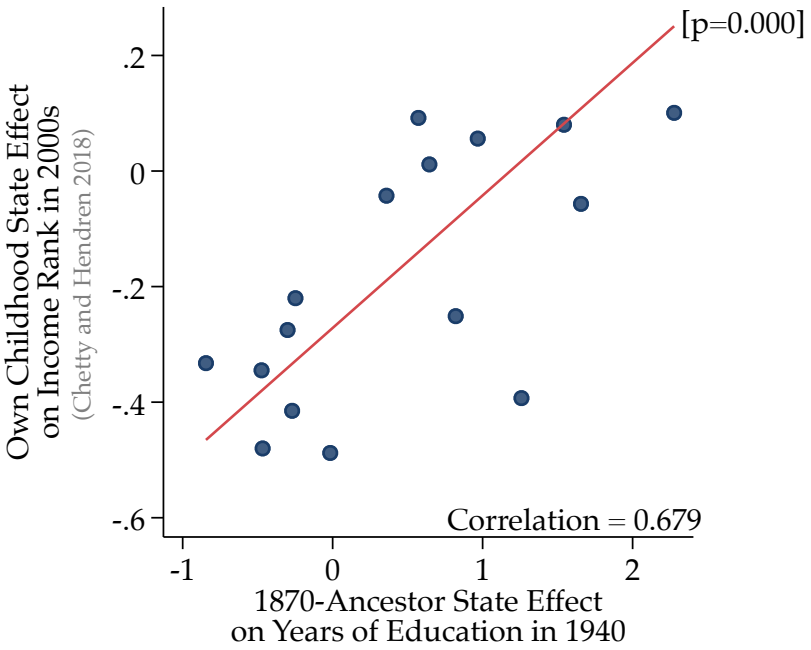
(E) Total Predicted Income



Notes: This figure shows the 1870 ancestor state of residence fixed effect estimates on years of education, literacy rate, homeownership rate, wage income, and predicted total income in 1940. A state's fixed effect is the deviation from the population-weighted average across all states after controlling for a quadratic function of age. The sample includes only Black prime-age men whose ancestors can be located in 1870.

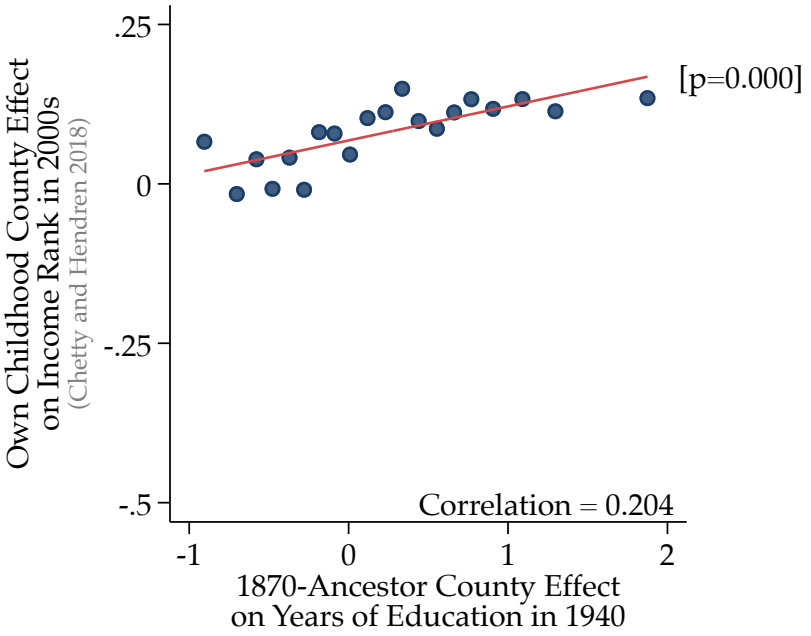
J.10 The Persistence of Place Effects

FIGURE J.29: Persistence of a State's Capacity to Generate Upward Mobility



Notes: This figure repeats Figure 16 after aggregating the county level estimates to the state level as an average weighted by population in 2000. See Data Appendix A for details on the sample and data.

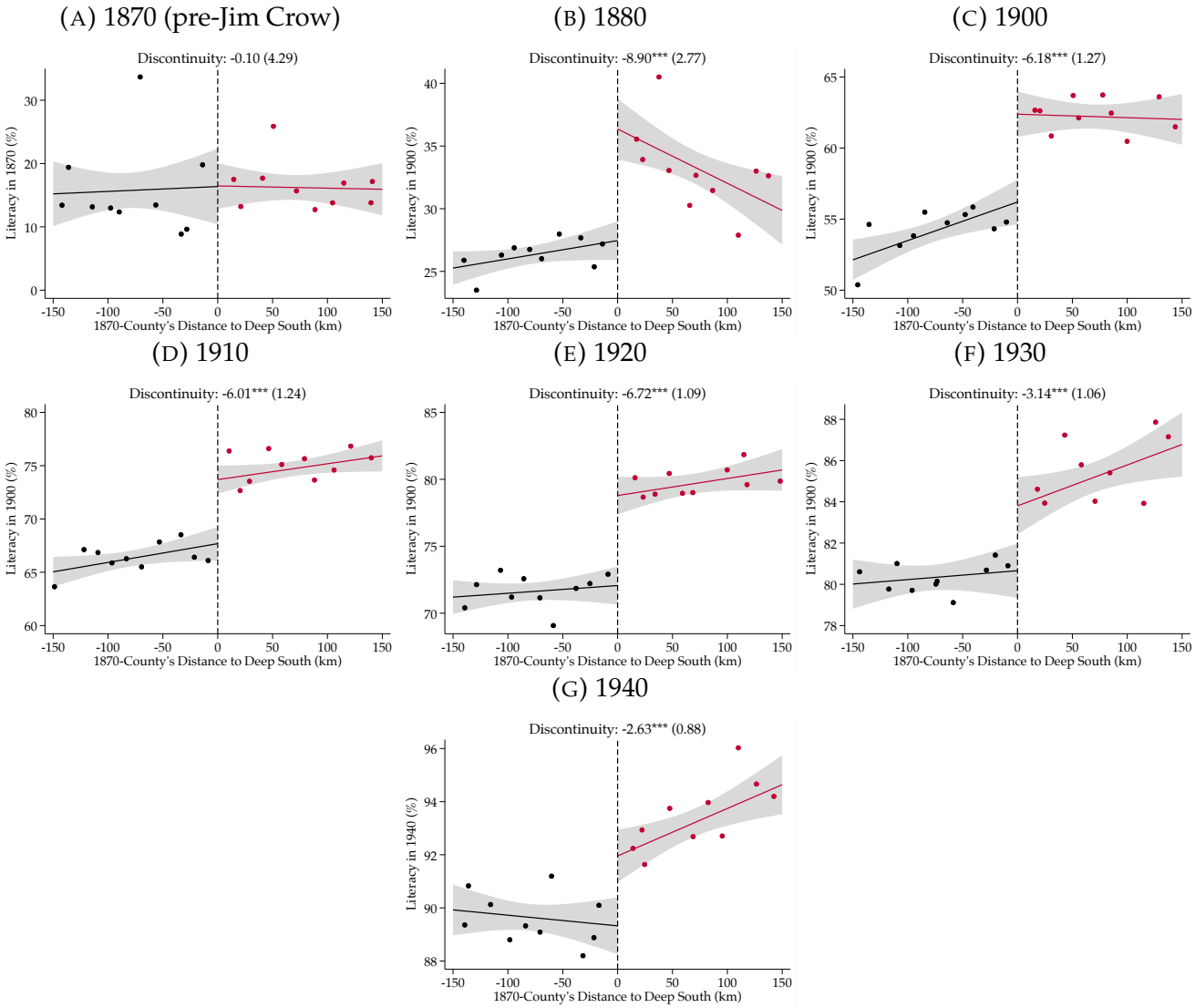
FIGURE J.30: Persistence of a County's Capacity to Generate Upward Mobility



Notes: This figure repeats Figure 16 using estimates of a county's effect on intergenerational mobility conditional on a child growing up with parents at the 75th (rather than the 25th) percentile of the national household income distribution. See Data Appendix A for details on the sample and data.

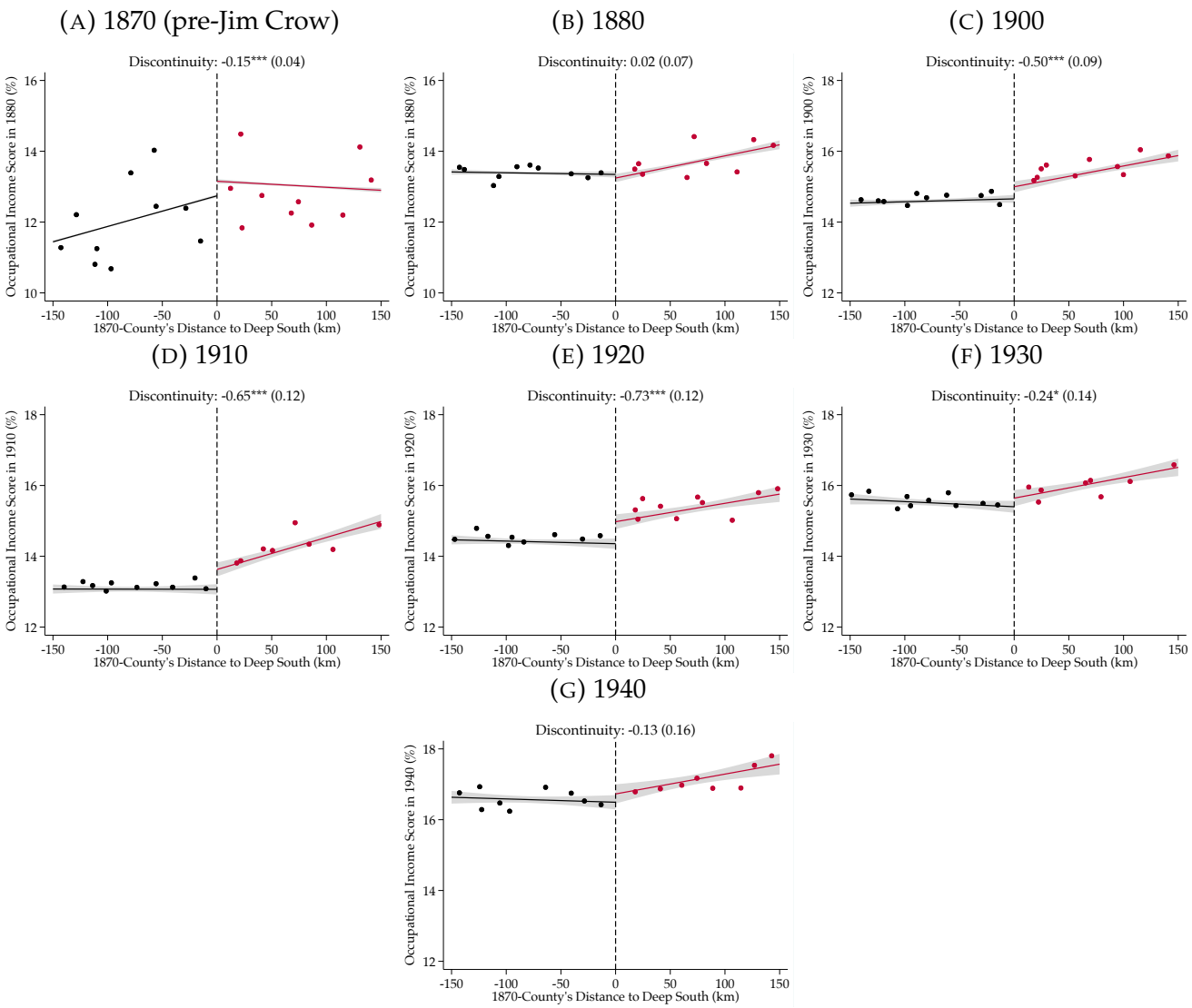
J.11 Additional Results from the Border Discontinuity Design

FIGURE J.31: Border Discontinuities in Literacy



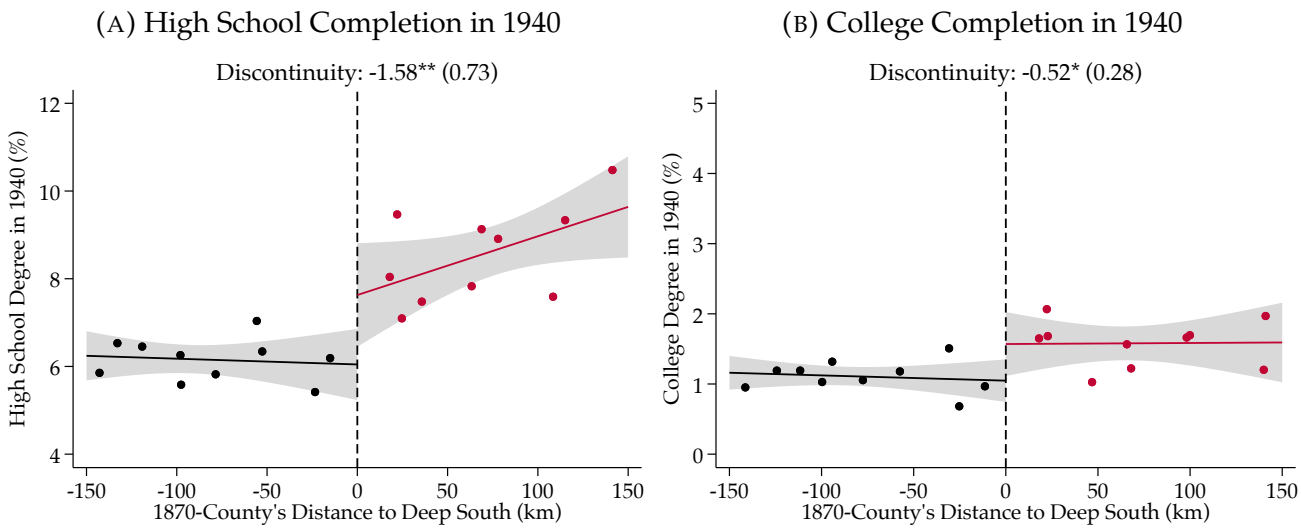
Notes: This figure shows binned scatter plots of counties' average literacy rate among formerly enslaved prime-age (20-54) men in 1860 by their ancestor's distance to the border of the Deep South in 1870. Lines represent the best linear fit of the individual-level data. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from the county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

FIGURE J.32: Border Discontinuities in Occupational Income Scores



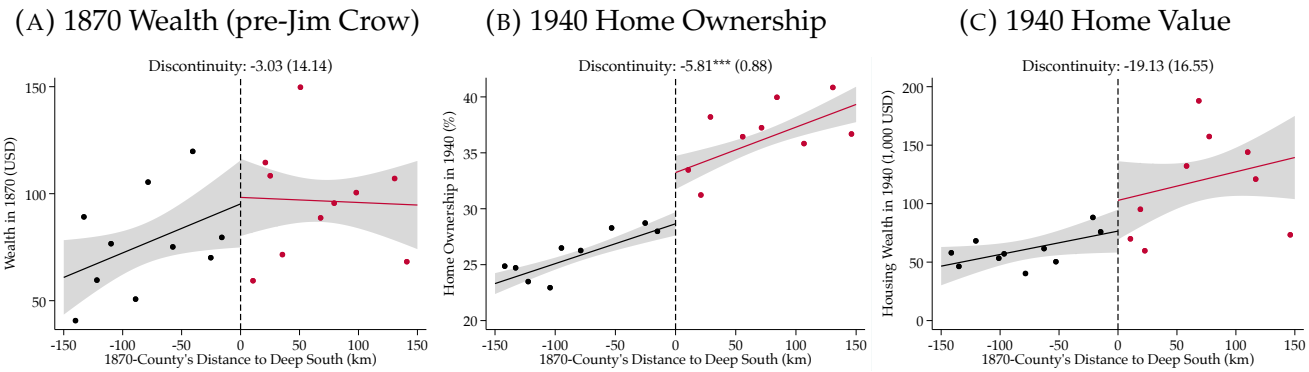
Notes: This figure shows binned scatter plots of counties' average occupational income scores among formerly enslaved prime-age (20-54) men in 1860 by their ancestor's distance to the border of the Deep South in 1870. Lines represent the best linear fit of the individual-level data. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from the county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

FIGURE J.33: Border Discontinuities in High School and College Completion



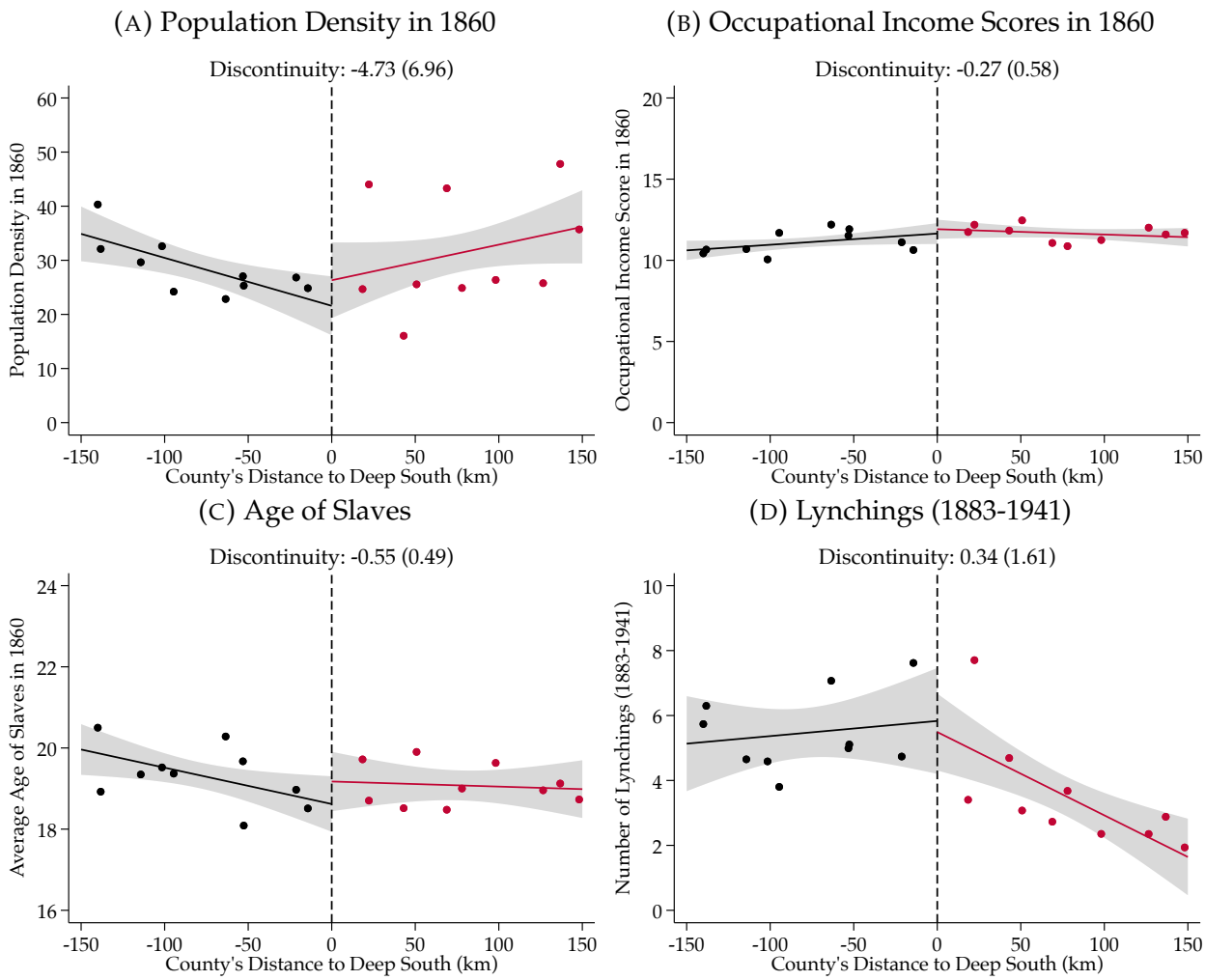
Notes: This figure shows binned scatter plots of counties' average high school and college completion rates among formerly enslaved prime-age (20-54) men in 1870 and 1940 by their ancestor's distance to the border of the Deep South in 1870. Lines represent the best linear fit of the individual-level data. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from the county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

FIGURE J.34: Border Discontinuities in Home Wealth



Notes: This figure shows binned scatter plots of counties' average wealth, home ownership, and house value (conditional on ownership) among formerly enslaved prime-age (20-54) men in 1870 and 1940 by their ancestor's distance to the border of the Deep South in 1870. Lines represent the best linear fit of the individual-level data. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from the county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

FIGURE J.35: No Border Discontinuities in Location Characteristics During Slavery

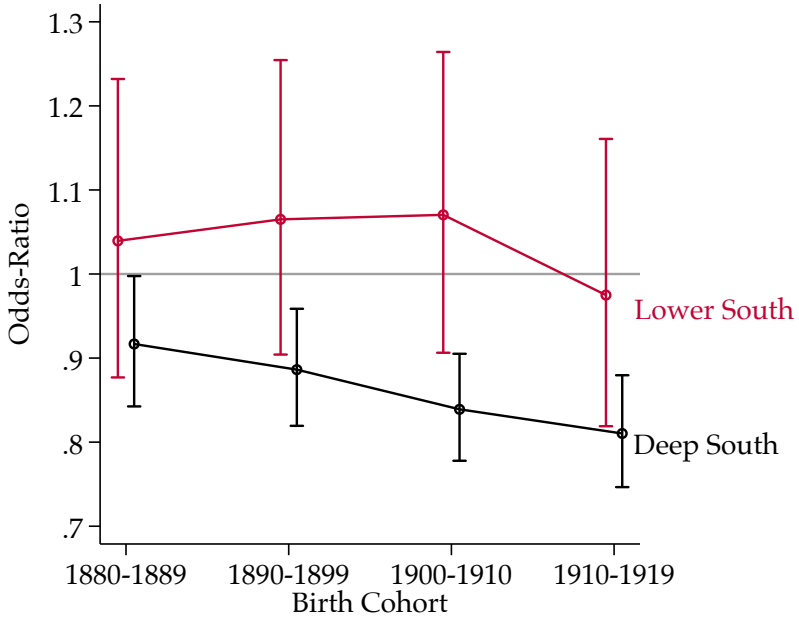


Notes: This figure shows binned scatter plots of counties' characteristics in 1860 or later by their distance to the border of the Deep South. Lines represent the best linear fit of the individual-level data. Counties inside (outside) the Deep South are assigned negative (positive) distances and are shown in black (red). Distances are measured from a county's centroid to the Deep South border. We limit our analysis to counties within 150 kilometers (93 miles) of the Deep South border. See Data Appendix A for details on the sample and data.

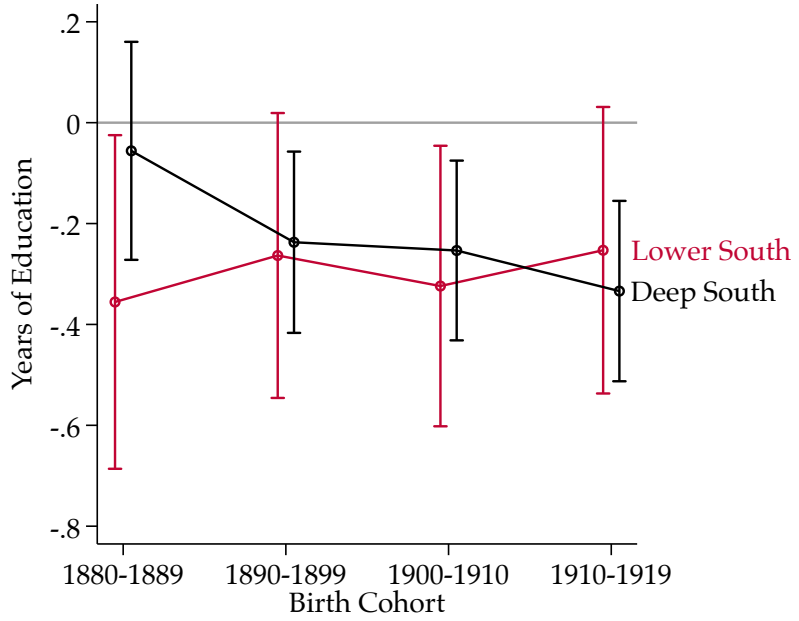
J.12 Cohort-Specific Effects of Exposure to Regional Factors

FIGURE J.36: Causal Effect of Exposure to Regional Factors across Birth Cohorts

(A) Literacy in 1940



(B) Years of Education in 1940



Notes: This figure shows the causal intent-to-treat (ITT) effect of Black men's exposure to the Lower South (in red) and the Deep South (in black) on education for different birth cohorts. The omitted baseline region is the Upper South; the omitted baseline birth cohort is 1870-1879 (of which eight years coincide with the Reconstruction era). Panel (A) uses literacy as an outcome, Panel (B) uses years of education. There is a clear pattern of decline for the Deep South, but not for the Lower South. The panels include 95 percent confidence bands that are clustered at the family level. See Data Appendix A for details on the sample and data.