

Jim Crow and Black Economic Progress After Slavery*

Lukas Althoff[†] Hugo Reichardt[‡]

[Most recent version here]

January 21, 2024

Abstract

This paper studies the long-run effects of slavery and restrictive Jim Crow institutions on Black Americans' economic outcomes. We track individual-level census records of each Black family from 1850 to 1940, and extend our analysis to neighborhood-level outcomes in 2000 and surname-based outcomes in 2023. We show that Black families whose ancestors were enslaved until the Civil War have considerably lower education, income, and wealth than Black families whose ancestors were free before the Civil War. The disparities between the two groups have persisted, not because of slavery per se, but because most families enslaved until the Civil War lived in states with strict Jim Crow regimes after slavery ended. In a regression discontinuity design based on ancestors' enslavement locations, we show that Jim Crow institutions sharply reduced Black families' economic progress in the long run. Jim Crow's educational restrictions likely played a key role driving its negative effect.

*We thank Lawrence Katz, Nathan Nunn, and five anonymous referees for their constructive comments that greatly improved this paper. We thank Leah Boustan, Davide Cantoni, David Card, Raj Chetty, Ellora Derenoncourt, Jeremiah Dittmar, Jonathon Hazell, Richard Hornbeck, Allan Hsiao, Ethan Ilzetzki, Ilyana Kuziemko, Camille Landais, David Lee, Trevon Logan, Ben Moll, Suresh Naidu, Steve Redding, Ricardo Reis, Maarten de Ridder, Bryan Stuart, Chris Walters, Tianyi Wang, Zach Ward, Gavin Wright, and numerous seminar participants for their insightful comments. We thank Andrea Bernini for sharing data with us. We also thank three anonymous graduate students at UMichigan for their mock referee reports. Tre' McMillan, Cynthia Nwankwo, and Bracklinn Williams provided excellent research assistance. This work was supported by Princeton's Program for Research on Inequality and its Industrial Relations Section.

[†]Stanford Institute for Economic Policy Research, Stanford University. lalthoff@stanford.edu

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

1. INTRODUCTION

Black Americans have faced a long history of economic oppression in the United States. Throughout the country's early history, slavery was legal—until around 1800 in the North and until the end of the Civil War (1861–1865) in the South. However, slavery was not the end of institutionalized oppression. Soon after slavery ended, Southern states created racially oppressive regimes to limit the economic progress of newly freed Black families—an institution called Jim Crow. States' Jim Crow regimes instituted racial segregation, Black voter disenfranchisement, and restrictions to Black Americans' geographic mobility.¹ After almost 100 years, the Civil Rights legislation of the 1960s outlawed racial discrimination and ended Jim Crow, making it “one of the most significant legislative achievements in American history” ([U.S. Senate, 2019](#)).

This paper studies whether and to what extent Black families' historical exposure to slavery and Jim Crow continues to shape their economic status. In sum, we find that the economic status of Black families today depends strongly on their historical exposure to those institutions. Black families everywhere left slavery with little or no measured physical or human capital. However, we show that Black families' economic progress after slavery critically depended on the state where they were freed. Most families enslaved until the Civil War were freed in the southernmost states. Those states implemented the most strict Jim Crow regimes, leading to high exposure that disproportionately disadvantaged the newly freed families for almost a century after slavery. Our results suggest that the severity of Jim Crow regimes is the main reason for the large differences in economic outcomes that persist between families enslaved until the Civil War and those freed an average of 50 years earlier.² We highlight denying equal access to education as a critical factor that likely made Jim Crow detrimental to Black economic progress.

We overcome the challenge of measuring each individual family's historical exposure to slavery and Jim Crow by tracing their census records from 1850 to 1940 using automated record-linking methods ([Abramitzky et al., 2019](#)). We construct family histories by combining the ability to follow individuals across full-count censuses and information on family interrelationships within households. First, to measure a family's exposure to slavery, we leverage that the 1850 and 1860 censuses did not record enslaved people. Therefore, we argue that we can identify families freed before the Civil War as those having ancestors recorded in the 1850 or 1860 census; others are classified as enslaved until

¹Throughout this paper, we use the term “Jim Crow” to refer to state-level institutions that limited Black Americans' civil rights. Examples include school segregation; vagrancy laws; or poll taxes. Factors that transcended state borders—such as lynchings or employer discrimination—or less prevalent instances of local Jim Crow-like ordinances do not fall under our definition of Jim Crow.

²Using aggregate counts of the Black population starting in 1790 and assuming that free Black families' population growth equaled that of white families, we approximate that the average free Black family was freed 50 years before the Civil War—around 1815.

the Civil War.³ We validate this method by developing a new surname-based approach to determine how likely a family was to have been enslaved until the Civil War (Ager et al., 2021).⁴ Second, to measure a family's exposure to Jim Crow, we use our linked sample to observe where a family's ancestors were enslaved and where they lived during Jim Crow.⁵ We quantify a state's Jim Crow intensity using various proxies, including our newly developed Jim Crow Index. This composite measure encapsulates different aspects of states' legislative efforts to suppress Black economic progress.⁶

To analyze the long-run economic impact of historical discrimination, we focus initially on a main sample comprising Black prime-age men in the linked census panel from 1850 to 1940. To extend our analysis to the present day, we add two sources of data. First, we link the 1940 census records to Social Security mortality data, which encompass nearly all individuals who died between 1980 and 2007. This extension allows us to derive proxies for late-life economic status based on a person's last neighborhood. Second, we use our surname-based approach that enables us to explore the relationship between ancestors' enslavement status and economic outcomes, not only in the full census populations but also in real-time data from a major US credit bureau.⁷

While exposure to oppression under slavery and Jim Crow was correlated, the two institutions' different geographies allow us to disentangle their effects. As a result of the rapid southern expansion of the US plantation economy, the longer a family was enslaved, the more likely they were to be concentrated in the southernmost states. Those states would become the epicenter of Jim Crow. Jim Crow regimes varied drastically across state borders; in contrast, slavery was an institution that transcended Southern states. Therefore, families who had been enslaved close to each other sometimes began to experience drastically different institutions of racial oppression under Jim Crow. Sometimes complicating escape from discriminatory regimes, Jim Crow restrictions added another barrier to geographic mobility.

We proceed in three steps to assess and disentangle the long-run effects of a family's exposure to slavery and Jim Crow. First, we divide our sample into two groups and doc-

³Linking the historical records of women remains difficult, allowing us to follow only the paternal line of ancestry. We estimate that intermarriage between families freed before 1865 and families freed in 1865 likely attenuates our estimates of the socioeconomic gaps between them in 1940 by one-third.

⁴This approach leverages changes in the distribution of surnames in the census from 1850–1860 to 1870–1880—before and after the inclusion of newly freed Black families—assigning a probability of having been enslaved until 1865 to each surname (see Appendix Table C.17). For example, the surname “Freedman” did not exist in 1860, but many newly freed families chose it in 1865. In contrast, the surname “Du Bois” became ten times less frequent in the census after it included the formerly Enslaved in 1870.

⁵As a family's enslavement location, we use their ancestor's state of birth or county of residence as observed in the 1870 census. We only use this information for families who were enslaved until 1865.

⁶The index is the principal component of critical factors in states' anti-Black institutions, as identified by historical literature: (1) proportion of race-specific laws discriminating against Black Americans, (2) disenfranchisement devices, (3) percentage of the population voting in Presidential elections between 1900 and 1940, (4) support for the Southern Manifesto, (5) lengths of Black school terms, and (6) the year a minimum wage for teachers was introduced.

⁷Due to data-sharing agreements, we cannot disclose the name of the credit bureau.

ument socioeconomic gaps between them: Black families who had higher exposure to both slavery *and* Jim Crow because they were enslaved until the Civil War (“Enslaved”); and families who were less exposed to both institutions because they were free earlier (“Free”). Second, we assess the importance of state-specific factors—including Jim Crow regimes—by decomposing this “Free-Enslaved gap” into variation in Black economic progress within and across ancestor states.⁸ To do so, we leverage plausibly exogenous variation in ancestors’ enslavement locations. Last, we use a regression discontinuity design to isolate the effect of states’ Jim Crow regimes from other factors that vary across states, such as economic activity, culture, or climate. Specifically, we compare the socioeconomic outcomes of Black families freed across state borders with more or less stringent Jim Crow regimes.

Our first key result is that today, Black families enslaved until the Civil War continue to have considerably lower education, income, and wealth than Black families freed before the Civil War. These Free-Enslaved gaps are almost half as large as the corresponding Black-white gaps. While immediately after slavery, the Free-Enslaved gaps were even larger, their narrowing has been much slower than one would expect under standard rates of intergenerational mobility. To mitigate the impact of error in measuring ancestors’ enslavement status, we validate our results with an instrumental variable strategy combining our surname- and linking-based measures.

Our second key result is that state-specific factors drive the long-run persistence of the Free-Enslaved gap. First, gaps due to direct exposure to slavery itself dissipated by 1940. In 1870, five years after the end of slavery, the socioeconomic status of recently freed families was far below that of families freed earlier, even for individuals from the same state. By 1940, those large Free-Enslaved gaps vanished conditional on the state in which their ancestors lived during slavery. Second, families enslaved until the Civil War were concentrated in the states where Black Americans fared worse after slavery. The difference in the two groups’ geographic distribution fully explains the persistently lower socioeconomic status of families enslaved until the Civil War. In sum, state-specific factors, while compressing the socioeconomic status of Black Americans *within* states, led to pronounced disparities *across* states, thereby placing descendants of those enslaved until the Civil War at a disproportionate disadvantage. It is important to note that our Free-Enslaved comparison isolates the additional disadvantage faced by descendants of those enslaved until the Civil War, without capturing the broader disadvantages stemming from slavery that affect all Black Americans, regardless of when they gained freedom.

Our third key result is that Jim Crow institutions are a root cause behind the state-

⁸If the main reason for the long-run persistence of the Free-Enslaved gap were differential exposure to slavery, we would expect this gap to largely reflect within-ancestor state differences between families freed before versus during the Civil War. In contrast, if the Free-Enslaved gap were driven by differential exposure to the state-specific factors, we would expect the gap to largely reflect across-state differences between families, irrespective of when they were free.

specific factors that drive the Free-Enslaved gap's persistence. Black families freed in states with more oppressive Jim Crow regimes experienced significantly lower rates of economic progress. The gaps in economic progress arise sharply at state borders and increase with the difference in states' Jim Crow intensity. For example, consistent with Louisiana's Jim Crow regime being far stricter than Texas's, we find that families freed in Louisiana attained 1.2 fewer years of education by 1940 than families freed only a few miles away in Texas. The long-run border discontinuity estimates (capturing the effects of institutional factors) are virtually identical in magnitude to the overall long-run state effects (capturing the effects of both institutional and non-institutional factors). These findings implicate state-level Jim Crow regimes as a central factor shaping the geography of Black economic progress.

We validate this border discontinuity design by showing that 1) differences in the socioeconomic status of formerly enslaved people only arise with the beginning of Jim Crow (circa 1880), 2) those differences are increasing in the intensity of states' Jim Crow regimes, 3) before Jim Crow there are no border differences in economic, agricultural, or demographic characteristics, and 4) Jim Crow regimes did not negatively affect white families. Basing our design on ancestor location before 1865—rather than the current location—leaves little room for selection, given that enslaved people had no say in their place of residence. Both historical and new empirical evidence support our main identifying assumption that an enslaved person's birthplace is exogenous to future generations' potential economic outcomes. While a family's enslavement location is a strong indicator of their exposure to Jim Crow, it is worth noting that many families migrated despite Jim Crow's institutional barriers to mobility ([Roback, 1984](#); [Naidu, 2010](#)) and high migration cost ([Carrington et al., 1996](#)). We assess the role of migration in shaping location effects using a standard framework of random assignment with imperfect compliance.

We explore how Jim Crow regimes slowed Black economic progress using a newly compiled dataset on state-level Jim Crow laws. We first classify Jim Crow laws by topic and find that the largest number pertains to education. Education is the target of 227 laws—over one-quarter of all Jim Crow laws passed throughout the South. Those laws racially segregated schools, reduced educational resources allocated to Black children, shortened term lengths for Black schools, and prevented Black Americans from participating in the local bodies that governed education. Statements from leading historians of the period confirm that educational restrictions were likely a key factor in Jim Crow's negative impact on Black economic progress.

This paper makes several contributions. First, leveraging new methods to link families' data across generations ([Abramitzky et al., 2020](#)), we generate new evidence on the mechanisms behind institutions' persistent effects ([Acemoglu et al., 2002](#); [Dell, 2010](#); [Donaldson, 2018](#); [Dell and Olken, 2019](#)). Second, we design methods to identify descendants of enslaved people, uncovering important socioeconomic differences among Black

Americans based on ancestral enslavement status. By analyzing exposure to Jim Crow, we find that systemic discrimination—the higher exposure to ongoing discrimination *because of past discrimination* (Bohren et al., 2022)—is central to the enduring legacy of racial oppression in the US. This aligns with seminal works that emphasize rapid progress by Black families where conditions allowed (Du Bois, 1935; Woodward, 1955; Ransom and Sutch, 2001; Aaronson and Mazumder, 2011; Naidu, 2012; Wright, 2013). Last, despite the recognized impact of location on upward mobility, its underlying causal mechanisms remain unclear (Olivetti and Paserman, 2015; Chetty et al., 2014; Chetty and Hendren, 2018). Our results show that institutions can play a key role in shaping upward mobility.

2. HISTORICAL CONTEXT

This section provides historical context for the evolution of racially oppressive institutions in the US—from slavery to Jim Crow and beyond.

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 had existed on American soil since the country’s colonial origins (Sowell, 1978). The roots of the free Black population may trace back to 1619 when settlers in Virginia purchased the first 20 Black people. Little is known about their fate, but it is likely that some of them were treated as servants who had to work for a fixed term and gained freedom afterward (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 some enslaved people, the Revolutionary War (1775–1783) provided a road to freedom. Responding to a need for troops and laborers, both the British and American leadership promised freedom to enslaved people willing and able to serve. It is estimated that up to 100,000 enslaved people ran away from plantations to do so (Schama, 2006). After the war, many remained in the US as free persons. As a result, the free Black population in some states increased dramatically.

The Revolutionary War also brought 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 a 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.

In the South, the path to freedom was narrow, especially in the Lower South.⁹ All Southern states except North Carolina allowed masters to free (“manumit”) their enslaved people 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, most manumitters freed all their enslaved people 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 (11.1 percent). 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 mainly originated from refugees who fled from Saint-Domingue (now Haiti) and the purchase of Louisiana from France, which had a sizable free Black population. By 1860, 40,000 of the 2.5 million Black Americans in the Lower South were free (1.6 percent).

The legal and economic status of free Black Americans varied greatly across locations and over time before 1865 (Sowell, 1978). In most states, free Black Americans were deprived of the right to vote and to hold political office. However, their legally protected property rights were respected in most cases. With the limited freedom they enjoyed, some free Black families 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 considerably across the country and, perhaps surprisingly, tended to be better further South (Berlin, 1976). In the North, free Black families were concentrated in cities where they suffered from competition with and hostility from white laborers (Frazier, 1949). Most free Black families in the South lived in rural areas, working as farmhands and casual laborers (Berlin, 1974).

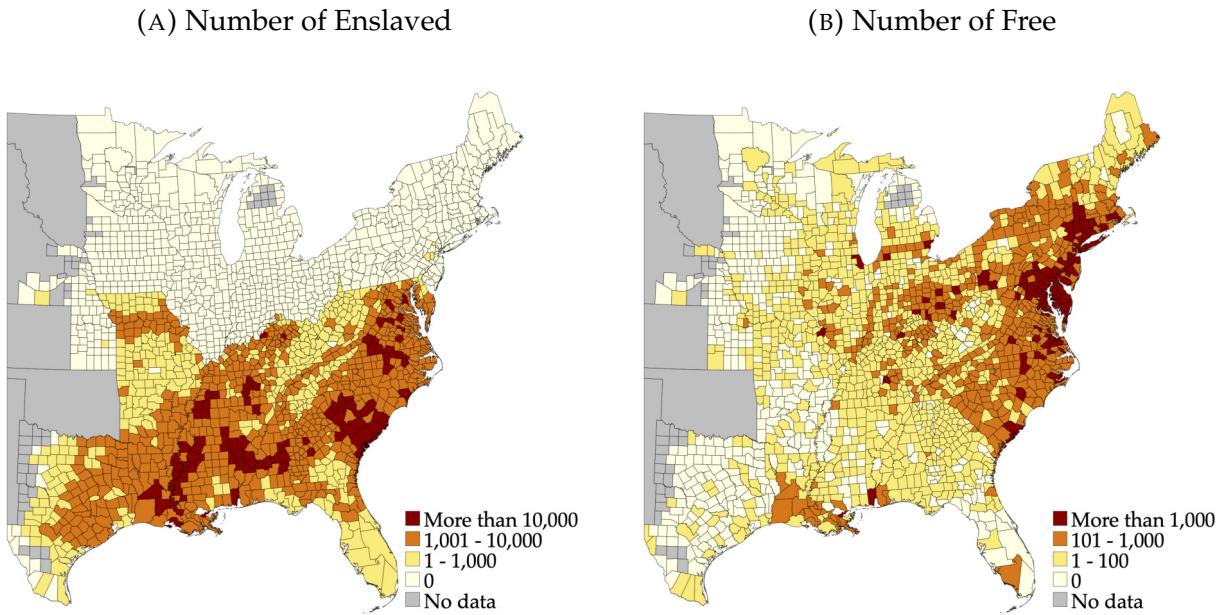
2.2 Freedom of All Black Americans after 1865

By the beginning of the Civil War (1861–1865), the enslaved population was concentrated in the Lower South (see Figure 1). The free Black population, in contrast, was concentrated in the North and the Upper South. These differences in geographic location exposed them to different institutional regimes after slavery.

The Civil War led to the emancipation of enslaved families, giving all Black Americans the same legal status. The average free Black family had likely already been free for around 50 years. For the first 12 years after the Civil War—the Reconstruction era (1865–1877)—the Union Army occupied the South. Black Americans experienced unprecedented economic progress under Reconstruction (Foner, 2014). New schools and

⁹The Lower South comprises Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, South Carolina, and Texas. The Upper South comprises Delaware, Washington, DC, Kentucky, Maryland, Missouri, North Carolina, Tennessee, Virginia, and West Virginia. The North comprises all other states.

FIGURE 1: Population by County in 1860



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. Appendix Figure C.39 shows the maps for 1790.

colleges 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](#)). Throughout Reconstruction, Black economic and political progress was met with violent opposition from white Southerners ([Du Bois, 1935](#); [Foner, 1963](#); [Blackmon, 2009](#)).

In 1877, the Union Army left the South, abandoning the project of Reconstruction. The disenfranchisement of Black people through legal and extra-legal means led to massive declines in Black political participation ([Kousser, 1974](#); [Naidu, 2012](#)). Many free Black Americans lost their higher social status and some left the South ([Woodson, 1918](#)).

Black Americans who remained in the South after Reconstruction faced increasing oppression through the rise of Jim Crow (1877–1964). Jim Crow regimes governed almost every aspect of Black life. Schools, workplaces, public transport, medical facilities, and parks were racially segregated ([Murray, 1950](#)). Poll taxes, literacy tests, and other rules limited Black suffrage ([Naidu, 2012](#); [Walton et al., 2012](#)). Enticement laws, contract enforcement laws, and emigrant-agent laws prevented Black workers from seeking economic opportunities with new employers or in states outside the South ([Roback, 1984](#); [Naidu, 2010](#)). Vagrancy laws criminalized the unemployment of Black people ([Blackmon, 2009](#)). In addition to legal factors, various extra-legal means of excluding Black Americans spread through the South and beyond.

From 1910 to 1940, many Black Americans started to leave the (Upper) South in the first wave of the Great Migration. Black families from the Lower South participated less

in this first wave, both because Jim Crow limited their geographic mobility and because migration was more costly for them (Roback, 1984; Naidu, 2010; Carrington et al., 1996). 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 (Boustan, 2016). However, opportunities in the North proved elusive to Black families (Akbar et al., 2020; Derenoncourt, 2022). In addition, even after the achievements of the 1960s, old forms of racial oppression have persisted, and new forms—such as mass incarceration and “color-blind” voter suppression—have arisen since (Western, 2006; Alexander, 2010; Bonilla-Silva, 2015; Darity et al., 2016). While significant progress has been made since the 1960s, including areas such as reducing residential segregation and improving professional representation, the reduction of racial disparities in crucial socioeconomic aspects, particularly men’s incomes, women’s family incomes, and overall wealth, has been limited (Cutler et al., 1999; Bayer and Charles, 2018; Althoff, 2021; Derenoncourt et al., 2022).

3. DATA AND NEW METHODS TO MEASURE A FAMILY’S EXPOSURE TO SLAVERY AND JIM CROW

A major empirical challenge we overcome in this paper is to measure a Black family’s exposure to slavery and Jim Crow. We construct family histories for Black Americans in the historical censuses and develop new methods to measure the two critical components of a family’s historical exposure to institutionalized oppression: how long a family was enslaved and where they were freed, determining the intensity of the Jim Crow regime under which they likely lived.

3.1 Measuring How Long a Family Was Enslaved

To measure how long a family was enslaved, we leverage that the pre-Civil War censuses of 1850 and 1860 did not record enslaved people.

Main method. We identify Black Americans free before 1865 (“the Free”) as those who were 1) recorded in the 1850 or 1860 census or 2) born in a state that had already abolished slavery; Black Americans who were born in slave states before 1865 and cannot be traced back to ancestors in the 1850 or 1860 census are classified as enslaved until 1865 (“the Enslaved”).¹⁰ We then carry this information forward to their descendants. To do

¹⁰We refer to Black families free before 1865 as “the Free” even though they or their ancestors may have been enslaved in previous decades. 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.

so, we build family trees using the census's information on family interrelationships for members of the same household and by linking individuals' records across time.

This classification strategy accurately identifies whether a Black family's ancestor was enslaved until 1865. In principle, if a family cannot be linked back to the 1850 or 1860 census, this could either mean that they were enslaved until 1865 or that they could not be linked using automated methods—for example, because their name was misspelled in one census. Hence, in the South, we inevitably misclassify some Black families who were free before 1865. However, census records show that only 6 percent of the Southern Black population were free in 1860. Many of these, we have accurately identified: record linkage in the 1870 census helped us identify around 20 percent of these individuals, and we found at least one male descendant in the 1940 census for approximately 10 percent. Therefore, our comparison involves a group almost certainly free in 1860 against a group where at least 94 percent were enslaved until the Civil War, minimizing the potential for attenuation bias due to imperfect linking rates (see also Appendix B.8).

Our classification method has two critical advantages over previous research, which typically relied on birthplaces to identify how long a family was likely enslaved. First, because the census only provides information on birthplaces for a person and their parents, the long-run effects of slavery cannot be studied in the census cross-section. Our panel allows us to follow families' individual-level outcomes until 1940 and the quality of their neighborhood until 2000. Second and most importantly, relying on a person's birthplace can only identify free Black families born in the North. However, 50 percent of all Black families free before 1865 lived in the South. Our method identifies a large number of those families. Measuring how long a family was enslaved and where it was freed is crucial to determining what role slavery, Jim Crow, and their interaction play in shaping the persistent effects of institutionalized racial oppression.¹¹

The Free-Enslaved gap quantifies disparities based on a family's male ancestry. Due to women's surname changes upon marriage, accurately linking female ancestry is challenging. Focusing on the male lineage minimizes bias that could arise from selective marriage patterns, allowing us to accurately estimate the Free-Enslaved gap as we define it. However, this approach limits our ability to estimate another important measure: the variation in socioeconomic status based on the proportion of Free vs. Enslaved ancestors across both maternal and paternal lines. Given the vast geographic and socioeconomic divides between Free and Enslaved families, intermarriage between these groups was likely limited by 1940. This is corroborated by quantitative evidence and historical narratives. However, we show that in the presence of intermarriage, even if minimal, the Free-Enslaved gap serves as a lower bound for the disparities between families with ex-

¹¹See Appendix Figure C.38 for average socioeconomic outcomes among descendants of the Enslaved and the Free by region of origin.

clusively versus no enslaved ancestors.¹²

Alternative method. We develop a second strategy to identify descendants of the Free and Enslaved based solely on surnames, without requiring census linkage. We use the change in the distribution over surnames from before 1865 (1850 and 1860 censuses), when the census included only free Black Americans, to after 1865 (1870 and 1880 censuses), when it included all Black Americans. This approach allows us to measure the likelihood that one's ancestors were enslaved until the Civil War among the full (not only the linked) sample of Black Americans in the census from 1870 to 1940 as well as real-time credit bureau data. The two approaches yield highly correlated Free-Enslaved classifications (see Appendix Figure C.37).

While some surnames were common among the Free and the Enslaved, others were characteristic of one group (see Appendix Table C.17). For example, the surname "Du Bois" was relatively frequent among free Black families in the 1860 census. However, with the inclusion of the families newly freed in 1865 in the 1870 census, Du Bois became ten times less frequent—an indication that having this surname meant a person likely descended from the Free. In contrast, the surname "Freedman" did not exist in the 1860 census but appeared in the 1870 census after some newly freed families chose it as their new surname. Thus, Black families called Freedman were likely enslaved until 1865.

3.2 Measuring the Exposure to State-Led Oppression During Jim Crow

Black families' exposure to slavery and Jim Crow is highly correlated. Families enslaved until 1865 were also geographically concentrated in states that would become the epicenter of Jim Crow. In contrast, families freed earlier were concentrated in states that would adopt less intensive Jim Crow regimes. These different geographic distributions result from the rapid southern expansion of the US plantation economy. The longer a family was enslaved, the more likely they were to be freed in the Lower South.

To measure a family's likely exposure to Jim Crow, we use that record linkage allows us to observe the birthplace of their formerly enslaved ancestors. While a family's enslavement location is generally a strong indicator of their exposure to Jim Crow over the subsequent 75 years, it is worth noting that many families migrated despite Jim Crow's institutional barriers to mobility (Roback, 1984; Naidu, 2010) and high migration cost (Carrington et al., 1996). Black Americans whose ancestors were enslaved in the Lower South were also very likely exposed to the Lower Southern states' strict Jim Crow regimes until 1940. Appendix Figure C.42 shows that prior to 1930 and 1940, the share of Black families originating from the Lower South who migrated out of the region was less

¹²In Appendix B.7, we derive this result theoretically. We estimate that for the first generation born after 1865, the gaps between Black Americans whose ancestors only descend from Enslaved vs. free Black ancestors could be 15 percent larger than the Free-Enslaved gap.

than 10 percent—significantly lower than the mobility rates experienced by Black families from the Upper South. Among families enslaved until the Civil War, the propensity to migrate North was especially low compared to Black families free earlier. We formally assess the role of migration in shaping ancestor locations’ long-run effects using a standard framework of random assignment with imperfect compliance.

Our primary measure of state regimes’ intensity is a new composite index we term the “Jim Crow index.” This index is derived from five factors frequently referred to in the historical literature as reflections of Jim Crow regimes: (1) the anti-Black discriminatory share of a state’s laws specific to race; (2) a state’s number of disfranchisement devices; (3) the share of congressional delegates who signed the Southern Manifesto; (5) the Black-white disparity in schools’ term lengths; and (6) the year minimum pay for teacher was introduced—legislation central to narrowing the large wage penalty historically suffered by Black teachers ([Card et al., 2022](#); [Cascio and Lewis, 2022](#)). Our measure builds on the existing Historical Racial Regime (HRR) index ([Baker, 2022](#)) but focuses on institutional factors and the Jim Crow era specifically.

To validate our findings, we consider alternative Jim Crow intensity measures. The HRR index itself, based on factors including the 1860 population share enslaved and the 1930 Black sharecropper share, is highly correlated with our Jim Crow index ($\rho = 0.99$). Additionally, we consider a state’s total number of Jim Crow laws. We analyzed 700 Jim Crow laws, digitizing “States’ Laws on Race and Color,” which aimed to document all race-related state laws in 1950 ([Murray, 1950](#)). We categorized each law as discriminatory (Jim Crow) or not based on its content and context provided by the author. We also incorporated additional laws on employment and suffrage not extensively covered in the primary source. The number of Jim Crow laws correlates with our Jim Crow index ($\rho = 0.73$). Another measure we consider is a new composite index of Black school quality, calculated from teacher salaries, student-to-teacher ratios, and term lengths for Black children in 1940—sourced from ([Card and Krueger, 1992](#))—negatively correlating with our Jim Crow index ($\rho = -0.94$).

It is essential to acknowledge the challenge in quantifying the severity of Jim Crow regimes. These regimes employed both legal methods, like literacy tests, and extra-legal tactics, such as voter intimidation, to marginalize Black Americans economically and politically. As Woodward noted, “[t]here [was] more Jim Crowism practiced in the South than there [were] Jim Crow laws on the books” (p. 102 [Woodward, 1955](#)). While no single measure can fully capture this complexity, we argue that a collective analysis of our proposed measures offers valuable insights into the nature and extent of Jim Crow institutions in different states.

3.3 Linked 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 if the match is *unique* conditional on race. We use a method that allows for misspellings by matching names based on their phonetic sound (NYSIIS). Allowing for misspellings tends to be a more conservative approach. Specifically, this method treats phonetically similar names as equivalent, yet maintains the requirement for uniqueness in establishing a match. Our data show a marginally higher match rate for free Black men compared to formerly enslaved men (18.5 vs. 17.1 percent, respectively, from 1870 to 1880).¹³ Because women tend to change their surname upon marriage, only men can be linked over time.

The census also contains information on the relationship between individuals in the same household. By observing a person in their parents' household during child- or adulthood, we can build family trees based on this information. We transfer parental data, such as Free-Enslaved status and county of residence, to subsequent census records of the individual and their descendants. These family trees allow us to study the evolution of a family's social, economic, and geographic mobility across generations. We study individuals' outcomes in census records between 1870 and 1940 (from the first census to include all Black Americans to the most recent full-count census available). Our primary outcomes include education, income, and wealth. Over time, the census data provide increasingly rich information on those outcomes. Therefore, we focus particular attention on the 1940 census.

To extend our analysis to the 21st century, we link the 1940 census to administrative mortality records between 1988 and 2005 ([Goldstein et al., 2021](#)).¹⁴ These records contain the neighborhood of residence (nine-digit ZIP code) at the time of death for the near-universe of deceased American citizens. We use National Historical Geographic Information System (NHGIS) data on each neighborhood's distribution of education, income, and wealth by race to proxy for a person's socioeconomic status. In effect, this sample contains individuals born before 1940 and deceased between 1988 and 2005.

To extend our results even further, we combine our surname-based Free-Enslaved classification with real-time data on individuals' economic outcomes from one of the primary US credit bureaus. The credit bureau merged our probabilistic classification with

¹³To evaluate linking rates by Free-Enslaved status, we contrast Black Americans born in the North (Free) with those from the South (mostly Enslaved), rather than basing the Free-Enslaved status on linkability in earlier decades. The relatively lower linking rates for Southern-born Black Americans may stem partly from their larger population sizes, which decrease the likelihood of having unique names within their birth states.

¹⁴The linkage from 1940 to 2000 leverages automated methods based on a person's name, year of birth, and state of birth ([Abramitzky et al., 2020](#)), analogous to the linkage between 1850 and 1940.

their universe of credit reports before removing personally identifying information. The credit bureau allows us to subset the data by race, sex, and age. The credit bureau does not observe a person’s race directly and instead predicts it based on the person’s first and last name as well as their neighborhood (nine-digit ZIP code). We access this anonymous data through a secure server.

3.4 Sample

For our analysis, we focus on Black men aged 20 to 54. For two reasons, we also limit our main sample to individuals who can be linked to their ancestors in 1880 or earlier. First, to identify a family who gained freedom before 1865 in a state that had not abolished slavery, they 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 more easily (e.g., because they have unique names) with those who cannot. Second, this restriction excludes families who immigrated to the US after 1880 and potentially experienced very different sets of institutions prior to their arrival. Our results are not sensitive to this restriction.

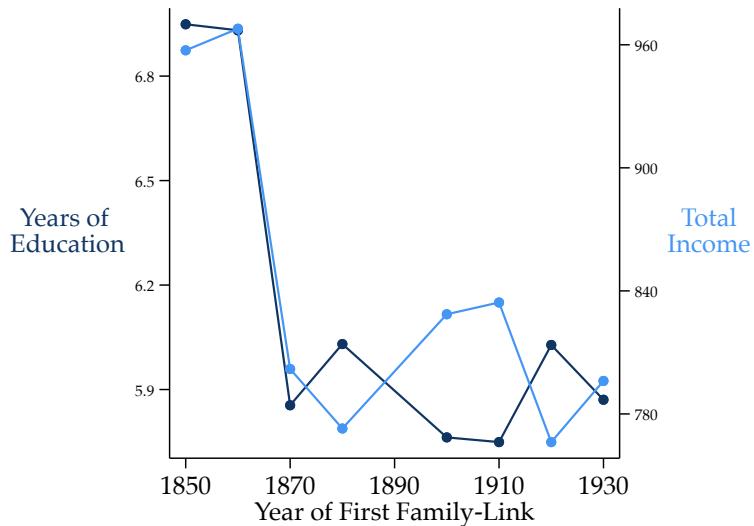
For 1940, our sample of Black prime-age men consists of 155,813 descendants of families enslaved until 1865 and 9,325 descendants of families freed before 1865. We link 10 percent of families in 1870 to at least one ancestor in 1940—an essential benchmark because those links allow us to observe the state in which a Black family’s ancestors were freed from slavery via their birthplace in the 1870 census. Imperfect linking rates are standard in the literature. Our family-level linking rates exceed those at the individual level (see Appendix Table C.14). Using our methodology, tracing an ancestor from 1870 to at least one descendant in 1940 is more feasible than following a person across multiple adjacent census waves for several reasons. First, to track an individual over time, we use links between both adjacent and non-adjacent census years. Second, we establish links between fathers and sons through their cohabitation, bypassing the need for census linkage. Third, the likelihood of establishing at least one link to a male descendant increases if an ancestor has multiple male descendants.

Our sample is highly balanced on observable characteristics (see Appendix Table C.15). For example, the literacy rate (20.4 percent) of those who we classify as formerly enslaved in our linked sample of 1870 matches the literacy rate of the 1870 Black census population—the vast majority of whom were enslaved until the Civil War. For free Black families in our linked sample of 1860, the literacy rate (65.1 percent) is also close to that of the 1860 Black census population (66.8 percent)—all of whom were free by definition of who was included in the census prior to 1865. The sample of individuals in 1940 linked to ancestors between 1850 and 1880 is also highly balanced compared to all Black men with US-born parents in 1940. From the 1940 census to administrative records in 2000,

we link 21,059 descendants of enslaved and 1,591 descendants of free Black families.

Potential Linking Bias. In constructing our main sample, we rely on linking families across census records. One may be concerned that linking procedures introduce mechanical differences between families enslaved until 1865 and those freed earlier. The most plausible concern is that a person's socioeconomic status depends on how many generations or decades they can be linked backward.

FIGURE 2: 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 dark blue line (left y-axis) shows the years of education; the light blue line (right y-axis) shows the total predicted income. The lines suggest no trend in outcomes outside of the break from 1860 to 1870. See Data Appendix C 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 2). In 1870, Black families enslaved until 1865 were included in the census for the first time. Consistent with that change in sample composition, we observe a significant drop in average income and education for 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 further do not have a mechanically higher socioeconomic status. 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.

4. A SIMPLE MODEL OF BLACK ECONOMIC PROGRESS AFTER SLAVERY

We propose a simple econometric model of Black economic progress to guide our interpretation of the forces that shape the Free-Enslaved gap's long-run persistence. Our framework incorporates intergenerational mobility, the effects of exposure to location-specific factors, (selective) migration, and the effect of delayed freedom. We use this model to answer the following questions: What factors determine the gap's long-run persistence? How important was the differential exposure to location-specific factors among the Enslaved and the Free in shaping the gap? Is the persisting disadvantage faced by descendants of the Enslaved a causal effect of slavery or Jim Crow?

4.1 Model setup

Let $y_{i,t}$ denote the human capital—or any other outcome of interest—for person 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 to which we can link families. We model $y_{i,t}$ to be determined by

$$y_{i,t} = \alpha_{i,t} + \gamma_{\ell(i,t)}^t + \rho y_{i,t-1} + \varepsilon_{i,t} \quad (1)$$

such that it depends on four factors: a factor capturing innate “ability” $\alpha_{i,t}$ with c.d.f. $F(\cdot)$, the family’s previous human capital $y_{i,t-1}$, their location $\ell(i, t) \in \mathcal{L}$, and a random error term $\varepsilon_{i,t}$ that satisfies $\mathbb{E}[\varepsilon_{i,t} | s_i, \alpha_{i,t}, \ell(i, t)] = 0$. Last, we define γ_{ℓ}^t as the effect of being exposed to location ℓ at time t . We model $y_{i,0}$ (the starting condition) as

$$y_{i,0} = \alpha_{i,0} + \gamma_{\ell(i,0)}^0 - \delta s_i + \varepsilon_{i,0}, \quad (2)$$

where s_i is an indicator for whether the family was enslaved until 1865. That is, in 1865, the outcomes depend on “ability,” location, and whether a person had been free before the Civil War. The parameter $\delta \geq 0$ captures any direct advantage that free Black Americans had relative to the Enslaved, such as access to education during slavery.¹⁵

¹⁵At time $t = 1$, the outcomes then become

$$y_{i,1} = (\lambda + \rho) \alpha_{i,0} + \rho \gamma_{\ell(i,0)}^0 + \gamma_{\ell(i,1)}^1 - s_i \rho \delta + \rho \varepsilon_{i,0} + \varepsilon_{i,1}, \quad (3)$$

where $\alpha_{i,1} = \lambda \alpha_{i,0}$ allows for transmission of “ability” over multiple generations. Thus, outcomes are determined by the “ability” of the initial generation through direct transmission of “ability” (λ) and through intergenerational advantage derived from “ability” in previous generations (ρ). The current location ($\gamma_{\ell(i,1)}^1$) shifts the level of a person’s human capital. Through intergenerational transmission, human capital is also affected by 1) how previous generations were impacted by where they lived ($\gamma_{\ell(i,0)}^0$), 2) whether their ancestors were enslaved until 1865 (δ), and 3) their ancestors’ idiosyncratic human capital shocks ($\varepsilon_{i,0}$).

4.2 The Intergenerational Effect of Being Enslaved Until the Civil War

We define the effect of descending from ancestors who were enslaved until the Civil War ($s_i = 1$) as the expected difference between the two groups in the absence of differences in “ability” ($\alpha_{i,0}$). That is, we define the average treatment effect as

$$ATE \equiv \int (\mathbb{E}[y_{i,1} | s_i = 1, \alpha_{i,0}] - \mathbb{E}[y_{i,1} | s_i = 0, \alpha_{i,0}]) dF(\alpha_{i,0}). \quad (4)$$

Throughout the paper, this definition will guide the interpretation of our estimates.

In conceptual contrast to prior work (e.g., [Sacerdote, 2005](#)), we argue that one should not think of slavery’s average treatment effect 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. Their enslavement status directly caused the location of enslavement, and the treatment effect should include its impact. From an econometric perspective, geographic location can be interpreted as a *bad control* since it is a mediating variable through which slave status affects future descendants ([Angrist and Pischke, 2008](#)).

5. SOCIOECONOMIC GAPS BETWEEN DESCENDANTS OF FREE AND ENSLAVED FAMILIES

This section documents the gaps in education, income, and wealth from 1870 to 2000 between descendants of families enslaved until the Civil War and those freed earlier. We find that these gaps are large and persist until today.

5.1 Evolution of the Free-Enslaved Gap until 1940

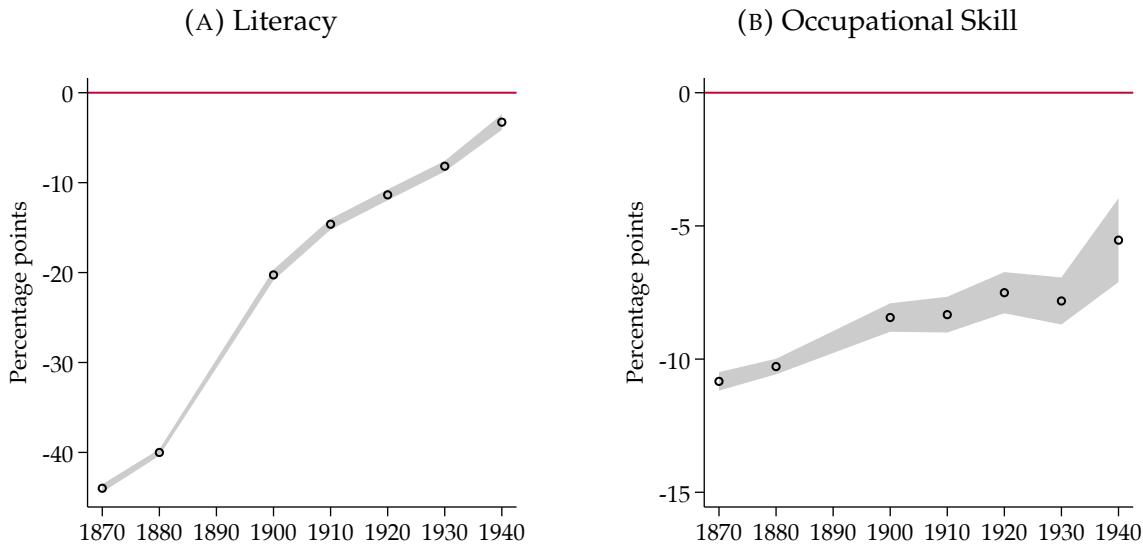
We estimate the Free-Enslaved gap (β_t) in socioeconomic outcomes ($y_{i,t}$) separately for each decade t from 1870 to 1940:

$$y_{i,t} = \alpha_t + \beta_t s_i + \phi_t' X_{i,t} + \varepsilon_{i,t}, \quad (5)$$

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 family level.¹⁶

¹⁶We define a family as a group of individuals with a common 1870 ancestor. In 1940, our linked sample comprises 49,876 families with an average of 1.6 prime-age male descendants each.

FIGURE 3: Free-Enslaved Gap (1870–1940)



Notes: This figure shows the gaps in literacy and occupation skill among prime-age (20-54) male descendants of enslaved vs. free Black Americans in each census decade. The sample includes both the South and North of the US. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. We assign “skilled” to occupations classified as “medium skilled workers” or above by the HISCLASS scheme ([Leeuwen and Maas, 2011](#)); and “unskilled” to others. 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. See Data Appendix C for details on the sample and data.

We find that the socioeconomic differences between descendants of the Free and Enslaved are large and persistent. In 1870, the formerly Enslaved were two times (over 40 percentage points) more likely to be illiterate than free Black Americans (see Figure 3). By 1940, the gap was 1.8 times (5 percentage points). Descendants of the Enslaved worked in less skill-intensive occupations than descendants of the Free from 1870 to 1940. Consistent with this skill gap, descendants of the Enslaved earn lower incomes and are significantly less likely to own their homes (see Appendix Figure A.2). Overall, we estimate the Free-Enslaved gap to be *smaller* than the gap between Black Americans born in the North vs. South before 1865—a comparison that [Sacerdote \(2005\)](#) uses as a proxy for the Free-Enslaved gap (see Appendix Figure A.1). Our estimates capture the important fact that free Black Americans fared far worse in the South than in the North after slavery.

The rich information on education, income, and wealth provided by the 1940 census allows us to get a detailed picture of the Free-Enslaved gap 75 years after slavery ended. Using those outcomes, we find that descendants of the Enslaved are less educated, earn lower incomes, and have accumulated less wealth than descendants of free Black Americans in 1940 (see Table 1). The gap in education amounts to 1.6 years—more than one-quarter of the average years of education among Black men in 1940. The likelihood that a descendant of the Enslaved earned a high school or college degree was only

half compared to descendants of the Free (see Appendix Table A.2). Consistent with the educational gap, the income and wealth of the Enslaved are lower.¹⁷

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 until Civil War	-1.59*** (0.05)	-145.92*** (6.13)	-7.24*** (0.62)	-694.69*** (65.85)
Controls (age, age ²)	Y	Y	Y	Y
% of Black-white gap	42	29	36	37
Adjusted R ²	0.04	0.05	0.01	0.01
Observations	163,549	154,463	164,357	46,971
<i>Ancestor Free</i>	9,078	8,551	9,070	3,227

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

The narrowing of the Free-Enslaved gap from 1870 to 1940 is slow relative to benchmark rates of intergenerational mobility among white Americans. To compare the convergence speed, we estimate socioeconomic gaps from 1870 to 1940 between white families whose ancestors had no measurable physical or human capital in 1870 and all other white families (see Appendix Figure B.29). In only 30 years, the gap in literacy between those two groups of white Americans rapidly shrunk from over 90 percentage points to less than 10 (from twice the Free-Enslaved gap in 1870 to half the Free-Enslaved gap in 1900). The homeownership gap for the two groups was similar to the respective Free-Enslaved gap in 1870 but closed by 1900—while the Free-Enslaved changed very little until then.

Robustness. We re-estimate the Free-Enslaved gap based on the *full* population (rather than the linked sample) of Black Americans in 1940 using our surname-based approach, yielding results very similar to our preferred approach based on record linking.¹⁸ The gaps between Black families holding surnames with high vs. low associated likelihoods of having been enslaved until the Civil War are $-1.40***$ (0.09) in years of education, $-113.15***$ (25.50) in wage income, $-2.31**$ (1.05) in homeownership, and $-1,098.68***$ (282.83) in house values.

¹⁷ Appendix Table A.1 compares the Free-Enslaved gap across different income measures.

¹⁸ See Appendix Table B.8. Without record linkage, we cannot assure that all Black families in the sample were present in the US during both slavery and Jim Crow. However, we can re-weight observations in the 1940 census to hold the distribution of surnames constant at its 1870 level.

Next, to mitigate misclassification bias, we use our surname-based measure as an instrumental variable (IV) for the linking-based measure. The resulting IV estimates offer an unbiased assessment of the Free-Enslaved gap if the errors in the linking-based measure are uncorrelated with the surname-based measure ([Ashenfelter and Krueger, 1994](#); [Angrist and Pischke, 2008](#)). This assumption is supported by the surname-based measure's independence from census-linking methods. These IV estimates suggest that measurement error reduces our initial estimates of the Free-Enslaved gap by an average of 9 percent across various outcomes (see Appendix B.8). For example, based on our IV estimates, descendants of the Enslaved attained 1.67*** (0.15) years less in education in 1940 than descendants of the Free, compared to 1.59*** (0.05) via OLS.

We also conduct an array of placebo exercises that validate our empirical strategy (see Appendix Tables B.11 and B.12). First, we use 1875 as a placebo year of emancipation. Specifically, we classify Black families as descending from the Free or the Enslaved based on whether or not we can link them back to ancestors in 1870 (rather than 1860). This placebo exercise yields no economically significant gaps. For example, a small gap of less than 1 percent in education emerges (compared to 25 percent in our baseline). Second, we use white Americans as a placebo group. Specifically, we divide white families into two groups depending on whether or not we can link them back to ancestors in the 1860 census, similar to our Free-Enslaved classification. Again, this placebo exercise yields no economically significant gaps. Last, adding various placebo groups as controls to our baseline specification leaves the overall patterns of persistence in the Free-Enslaved gap unaffected.

5.2 The Free-Enslaved Gap in the 21st Century

The Civil Rights Movement (1954–1968) ended Jim Crow, thereby instigating institutional change that held the promise to accelerate Black economic progress. Existing evidence indeed suggests that Black Americans' social mobility temporarily surged around 1970 ([Clark, 2014](#)). How has the Free-Enslaved gap evolved since the end of Jim Crow?

We extend our analysis past 1940 using two methods. First, we merge data from a major US credit bureau with our surname-based probabilities of descending from ancestors enslaved until the Civil War. This approach lets us estimate the Free-Enslaved gap in real-time without needing record linkage. We use a snapshot of this data from March 2023, limiting the main sample to Black Americans as identified by the credit bureau through names and nine-digit ZIP codes. Second, we link 1940 census records for Black Americans to administrative mortality data circa 2000, covering birth cohorts from 1910 to 1940. These records include a person's last residential nine-digit ZIP code, allowing us to infer neighborhood proxies for their income, wealth, and education. Our results are robust to excluding high-mortality neighborhoods (e.g., retirement communities).

TABLE 2: Free-Enslaved Gap (2023)

	Total income (USD) Mean: 92,068.48	Disposable income (USD) Mean: 52,773.74	Credit Score (from 300 to 850) Mean: 630.41	Hourly Job Mean: 0.72
Ancestor Enslaved until Civil War	-12,487.72*** (1,147.08)	-11,623.44*** (920.12)	-33.15*** (2.07)	0.05*** (0.01)
Controls (age group-FE)	Y	Y	Y	Y
% of Black-white gap	23	26	40	69
Adjusted R ²	0.001	0.001	0.003	0.000
Observations	547,189	547,189	547,189	459,889

Notes: This table shows the Free-Enslaved gap in total income, disposable income, credit score, and hourly-wage employment among Americans as of March 2023. We estimate a person's likelihood to descend from free Black Americans via their surname, not requiring record linkage. We re-weight the sample to hold the distribution of surnames constant at the 1870 level. The sample's average likelihood of a person's ancestor to be free before the Civil War based on their surname is 9.6 percent—very close to the factual fraction. The sample includes both the South and North of the US. Credit scores (VantageScore® 3.0) reflect a person's credit health, ranging from 300 to 850 (scores above 700 are considered "good" and scores below 550 "very poor"). See Data Appendix C for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Using US credit bureau data from 2023, we find that descendants of the Enslaved have vastly lower incomes and worse credit health than descendants of free Black Americans (see Table 2). The Free-Enslaved gap in disposable incomes is \$11,620 (22 percent of Black Americans' average). The Free-Enslaved gap in credit scores is 33 points (one-fifth of the difference between "good" and "very poor" credit). Descendants of the Enslaved are also more likely to work in hourly-wage jobs, presumably leading to higher uncertainty in earnings compared to salaried jobs. These Free-Enslaved gaps amount to 23 to 69 percent of the corresponding Black-white gaps.

Using neighborhood-level data from mortality records linked to the 1940 census, we find that around 2000, descendants of the Enslaved resided in neighborhoods with lower education, income, and wealth than those of the Free descendants (see Appendix Table A.3). Descendants of the Enslaved lived in neighborhoods where Black residents were 3.9 percentage points less likely to hold a high school degree and 2.6 percentage points less likely to hold a college degree. Black residents' expected incomes in those neighborhoods were \$5,100 lower (17 percent of the median). Owner-occupied houses were worth \$17,500 less (19 percent of the median). Because those estimates ignore within-neighborhood differences, they should be considered an underestimate of the actual Free-Enslaved gap.¹⁹

¹⁹The distribution of deaths across space in our data highly correlates with population density ($\rho = 0.91$). Our results are also robust to dropping neighborhoods with excess mortality (see Appendix Table B.9).

In sum, our two strategies suggest that the present-day Free-Enslaved gaps in various economic outcomes amount to at least one-fifth of the corresponding Black-white gaps. This finding highlights the enduring impact of historical oppression on present racial disparities. Importantly, the Free-Enslaved gap only quantifies the *additional* disadvantage faced by those whose ancestors were enslaved until 1865 compared to those who gained freedom earlier. Most Black families, even those who were free before the Civil War, were enslaved in earlier periods, and all Black Americans faced discrimination due to slavery and Jim Crow, regardless of their specific family history. The sheer difference in intensity of their experiences yields socioeconomic gaps of such enormous magnitude. Next, we turn to the drivers of this persistence.

5.3 Interpreting the Free-Enslaved Gap

Using our model from Section 4, the Free-Enslaved gap measured as $\hat{\beta}_{1940}$ in equation (5), is a consistent estimator of

$$\begin{aligned} \mathbb{E}[y_{i,1} | s_i = 1, X_{i,t}] - \mathbb{E}[y_{i,1} | s_i = 0, X_{i,t}] &= \\ (\lambda + \rho) (\mathbb{E}[\alpha_{i,0} | s_i = 1, X_{i,t}] - \mathbb{E}[\alpha_{i,0} | s_i = 0, X_{i,t}]) + \\ \mathbb{E}[\rho\gamma_{\ell(i,0)}^0 + \gamma_{\ell(i,1)}^1 | s_i = 1, X_{i,t}] - \mathbb{E}[\rho\gamma_{\ell(i,0)}^0 + \gamma_{\ell(i,1)}^1 | s_i = 0, X_{i,t}] - \rho\delta. \end{aligned}$$

Intuitively, the Free-Enslaved gap, therefore, reflects 1) any potential differences in “ability” between the two groups transmitted over generations, 2) different exposure to locations over time (as a result of slavery and potential selection), and 3) the inherited disadvantage of descending from an enslaved person conditional on environment and “ability.” In the next section, we show that the two groups’ differential exposure to locations due to slavery—not selection—accounts for virtually all of the Free-Enslaved gap.

6. THE IMPORTANCE OF GEOGRAPHY IN SHAPING BLACK ECONOMIC PROGRESS AFTER SLAVERY

In this section, we use ancestors’ enslavement locations as plausibly exogenous variation in where Black families lived to identify what fraction of the Free-Enslaved gap is caused by differential exposure to place-specific factors. Our main sample is limited to descendants of ancestors enslaved until the Civil War, thereby narrowing the sample to the South. We find that state-specific factors are the leading cause of the Free-Enslaved gap’s persistence after 1940.

6.1 States' Effect on Black Economic Progress After Slavery

We estimate each state's causal effect on the long-run economic progress of Black families freed there in 1865 (excluding free Black Americans and their descendants). Our empirical strategy to identify the importance of exposure to location-specific factors builds on the following assumption, which we discuss in detail in Section 6.4.

Assumption 1 (Exogeneity of enslavement location). *The enslaved population was not selected into location. That is,*

$$\alpha_{i,0} \perp\!\!\!\perp \ell(i,0) \text{ if } s_i = 1$$

where s_i is a dummy variable equal to 1 if one's ancestor was enslaved up to 1865, $\ell(i,0)$ is the birthplace of one's enslaved ancestor, and $\alpha_{i,0}$ is the innate "ability" of one's enslaved ancestor.

We limit our sample to families whose ancestors were enslaved until the Civil War and estimate the causal effect that the geographic distribution of formerly enslaved ancestors had on the Black economic progress of their descendants:

$$y_i = \eta_{\ell(i,1865)} + \phi' X_i + \epsilon_i, \quad (6)$$

where y_i are socioeconomic outcomes in 1940 and X_i is a vector of controls as defined in equation (5). In the context of the model introduced in Section 4,

$$\eta_\ell = \rho \gamma_\ell^0 + \mathbb{E}[\gamma_{\ell(i,1)}^1 \mid s_i = 1, \ell(i,0) = \ell, X_i], \quad (7)$$

where γ_ℓ^0 and γ_ℓ^1 are the effects that location ℓ had on Black families during and after slavery respectively. Thus, η_ℓ reflects both the (inherited) effect the state of birth ℓ had on the ancestor during slavery and the expected effects of future locations of their descendants given the 1865 location. One can interpret η_ℓ as an intent-to-treat (ITT) effect of living in location ℓ from before the Civil War to 1940, where the initial location is plausibly randomly assigned, but the post-1865 location is a result of endogenous (and potentially selective) migration decisions.

The effect of being freed in each state in 1865. We find a distinct geography of Black economic progress after slavery (see Appendix Figure A.6). Gaining freedom in a state further south negatively affected Black families' socioeconomic outcomes in the long run. For example, a family freed in Louisiana would attain over two years more education had they instead been freed in, say, Kentucky.²⁰ States affect other outcomes, such as literacy, income, or wealth, with similar magnitudes.²¹ States' effects are substantial even in 2000

²⁰Being freed in Louisiana has the strongest negative impact on education by 1940 (-0.84 years less than the average across Southern Black Americans)—followed by Georgia and South Carolina (-0.47 years). Missouri has the strongest positive impact (2.28 years), followed by Kentucky (1.66 years).

²¹Note that the 1940 census lacks information on non-wage income; wage income alone tends to be

when, for example, families freed in Louisiana have average incomes lower by over one-quarter of the average income among Black Americans compared to those rooted in states along the border to the North.

The effect of living in each state between 1865 and 1940. We formally assess the importance of post-slavery migration and recover the effect of living in each location ℓ between 1865 and 1940 on Black economic progress absent migration (γ_ℓ^1). We do so based on Assumption 1 and the additional assumption that place-specific experiences during slavery ceased to affect descendants in 1940 directly ($\rho\gamma_\ell^0 = 0$); we formalize this decomposition in Appendix B.9. This problem is a standard case of multiple instruments (location assignment) and imperfect compliance (migration). Specifically, the intent-to-treat effect of initial location ℓ , η_ℓ , is the average of all potential future locations' treatment effects, $\gamma_{\ell'}^1$, weighted by the probability of migrating from ℓ to ℓ' :

$$\eta_\ell = \sum_{\ell' \in \mathcal{L}} p_{\ell,\ell'} \cdot \gamma_{\ell'}^1.$$

We invert the migration probability matrix to recover the effect of living in each state until 1940, which is unaffected by selective migration under the assumption that the average innate “ability” of Black Americans in 1865 did not differ across enslavement locations.

Our results indicate that the effect of being freed in location ℓ closely approximates the treatment effect of living in ℓ from 1865 to 1940 (see Appendix B.9). The recovered treatment effects are almost identical to the intent-to-treat effects estimated using equation (6), except for the border states of the Upper South. The effect of living in the border states is more negative than the effect of being freed there, suggesting that the relatively better conditions for Black Americans were partly due to greater migration opportunities. For those freed in the Lower South, benefits from Northern opportunities were more limited due to lower migration rates and a reduced likelihood of the North being their destination conditional on migration.

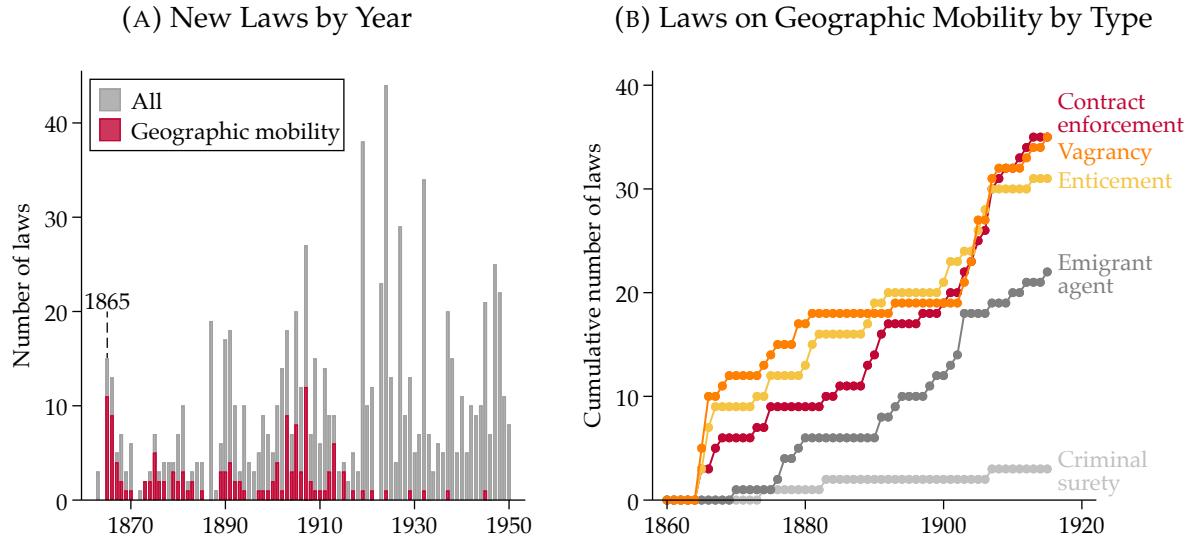
With Black families freed in the Lower South faring so much worse than those freed elsewhere, it may seem puzzling why the region did not experience a larger exodus than the Upper South. For example, 75 percent of Black families enslaved in Louisiana still lived there in 1940; less than 10 percent reached the North (see Appendix Figures C.40 and C.41). Lower Southern white families were almost 30% more likely to migrate. Institutional and economic factors partly resolve this puzzle.

First, Jim Crow directly targeted the geographic mobility of Black people (Roback, 1984; Cohen, 1991; Naidu, 2010): enticement laws and contract enforcement laws limited Black workers' ability to terminate their employment contracts; vagrancy laws criminalized being out of employment; emigrant-agent laws prevented employers from seeking

uninformative for farmers and other self-employed occupations. Total income measured in 2000, however, yields estimates that line up well with our evidence on 1940 outcomes in education and wealth.

workers from other states; criminal surety laws created the possibility of involuntary servitude upon arrests for minor charges (see also [Blackmon, 2009](#)). These laws began emerging immediately after slavery (see Figure 4).

FIGURE 4: Number of Jim Crow Laws Across the South



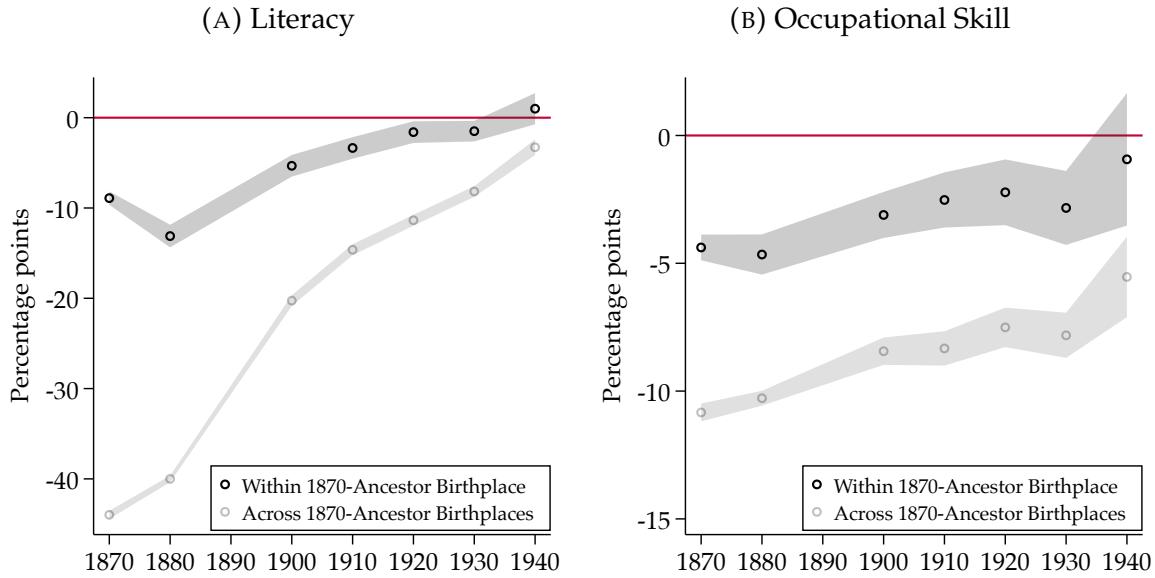
Notes: This figure shows the number of new Jim Crow laws passed across all Southern states each year (Panel A) and the cumulative number of laws pertaining to the geographic mobility or employment of Black Americans by type (Panel B). See Data Appendix C for details on the data.

Second, moving to the North was costly, especially from the Lower South. Among families enslaved until the Civil War, the propensity to migrate North was especially low compared to Black families free earlier—some of whom may have used the resources they had accumulated by the end of the Civil War to leave the South. The region’s geographic distance to the North limited the potential of social networks to lower the cost of migration ([Carrington et al., 1996](#)). Moreover, despite successful migration to the North, many Black families still faced challenges in capitalizing on available opportunities ([Collins, 1997](#); [Akbar et al., 2020](#); [Derenoncourt, 2022](#)).

6.2 The Free-Enslaved Gap is Driven by Geography

To explore the importance of differential exposure of state-specific factors, we first compute the Free-Enslaved gap conditional on ancestor location. To do so, we add fixed effects for the state of birth ℓ of a family’s ancestor before 1865 to our baseline specification in equation (5). This exercise provides a back-of-the-envelope assessment of how important geography was in shaping the Free-Enslaved gap’s long-run persistence. It does not account for free Black Americans’ potential selection into states before 1865. In the next section, we account for this potential selection and provide a lower bound for the causal importance of state-specific factors.

FIGURE 5: Free-Enslaved Gap Conditional on Ancestor State (1870-1940)



Notes: This figure shows the gaps in literacy and occupational skill before (light) and after (dark) including fixed effects for 1870 ancestor state of birth. The sample includes both the South and North of the US. The comparison is made between prime-age (20-54 years) male descendants of enslaved vs. free Black Americans in each census decade. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. We assign “skilled” to occupations classified as “medium skilled workers” or above by the HISCLASS scheme ([Leeuwen and Maas, 2011](#)); and “unskilled” to others. Both panels control for age and include 95 percent confidence bands clustered at the family level. See Data Appendix C for details on the sample and data.

We find that in contrast to the unconditional Free-Enslaved gap, the conditional gap was large in 1870 but shrunk to virtually zero after 1940 (see Figure 5).²² The 1940 gap in literacy, for example, fully closes after accounting for variation across ancestor states. Similarly, the conditional Free-Enslaved gap in 2000 is insignificant for all outcomes (see Appendix Table A.6). These results suggest that the Free-Enslaved gap persists solely because the two groups were exposed to different state-specific factors after slavery.

It is ancestor *states* that explain the Free-Enslaved gap, not other levels of ancestor geography (see Appendix Figure A.4). We find that the gap conditional on ancestor *region* is still large after 1940, suggesting that the Free-Enslaved gap is not merely a result of North-South differences. Adding ancestor *county* fixed effects does not further explain the Free-Enslaved gap, suggesting that it is not geographic granularity that makes states an important explanation.

With the ancestor state accounting for the vast majority of the Free-Enslaved gap, there is little room for other factors—such as differences in “ability” or the advantage of being free earlier—to drive the gap after 1940. State-specific factors compressed the socioeconomic status of Black Americans within states irrespective of their ancestors’ enslavement.

²²The 1940 gaps in almost any other outcome also shrink to zero after conditioning on the 1870 state of origin (see Appendix Figure A.3 and Appendix Table A.5).

ment status (see Appendix Figure A.5). Their high exposure to states that slowed Black economic progress after slavery placed descendants of the Enslaved at a disproportionate disadvantage.

Two exercises provide additional evidence against the importance of alternative explanations for the Free-Enslaved gap's persistence. First, we consider free Black Americans who had no measured physical or human capital by the end of slavery. We find that even this group of free Black Americans had higher socioeconomic status than descendants of the Enslaved by 1940 (see Appendix Table A.4). This result supports the conclusion that the Free-Enslaved gap's persistence is not driven by selection into freedom or the advantage of being free earlier. Second, we estimate the Free-Enslaved gap controlling for skin tones. We find that the Free-Enslaved gap is almost identical with or without this control.²³ This result rules out potential differences in discrimination of descendants of the Free and the Enslaved based on their skin tones driving the gap's persistence (see also Abramitzky et al., 2023).

6.3 Lower Bound for the Causal Effect of Geography

We develop a lower bound for how much of the Free-Enslaved gap is caused by the two groups' differential exposure to state-specific factors. Specifically, we can decompose the average treatment effect (ATE) of descending from ancestors enslaved until the Civil War into the sum of 1) the inherited disadvantage conditional on location and "ability" (δ), 2) the geographic effect of the ancestor's enslavement location (*geographic endowment effect*), and 3) the effect on the ability to migrate to more favorable locations (*location choice effect*). Formally, we decompose the ATE as defined in Section 4.2 as

$$ATE = -\rho\delta + \theta + \kappa$$

where θ is the *geographic endowment effect* and κ is the *location choice effect*.²⁴

We argue that the geographic disadvantage that the Enslaved population faced relative to the Free *within the South* provides a lower bound (in absolute terms) for the *geographic endowment effect*. In the North, descendants of the Free tended to face more favorable conditions after slavery than those in the South. A large part of the *geographic endowment effect* therefore likely results from the fact that around half of the Free popula-

²³See Appendix Figure B.18 for results controlling for a person's classification as "Mulatto" or "Black."

²⁴Imposing Assumption 1,

$$\theta \equiv \int \sum_{\ell \in \mathcal{L}} \left(\Pr(\ell_{(i,0)} = \ell | s_i = 1) - \Pr(\ell_{(i,0)} = \ell | s_i = 0, \alpha_{i,0}) \right) \times \left(\rho \gamma_\ell^0 + \mathbb{E} [\gamma_{\ell(i,1)}^1 | s_i = 1, a_{i,0}, \ell_{(i,0)} = \ell] \right) dF(\alpha_{i,0}) \quad (8)$$

$$\kappa \equiv \int \sum_{\ell \in \mathcal{L}} \Pr(\ell_{(i,0)} = \ell | s_i = 0, \alpha_{i,0}) \times \left(\mathbb{E} [\gamma_{\ell(i,1)}^1 | s_i = 1, a_{i,0}, \ell_{(i,0)} = \ell] - \mathbb{E} [\gamma_{\ell(i,1)}^1 | s_i = 0, a_{i,0}, \ell_{(i,0)} = \ell] \right) dF(\alpha_{i,0}). \quad (9)$$

tion lived in the North before 1865—an effect that we ignore to provide a lower bound. Formally, we assume that *geographic endowment effect* $\leq Z$ where Z is defined as

$$Z \equiv \sum_{\ell \in \mathcal{L}} \left(\Pr(\ell_{(i,0)} = \ell | s_i = 1) - \Pr(\ell_{(i,0)} = \ell | s_i = 0, \ell \in S) \right) (\eta_\ell - \eta_{\ell'}) \quad (10)$$

where $S \subset \mathcal{L}$ denotes all states in the South, $\ell' \in S$ is an arbitrary reference state in the South, and $\eta_\ell - \eta_{\ell'}$ as defined in (7) is the intent-to-treat effect of having a formerly enslaved ancestor born in state ℓ (relative to state ℓ'). The *geographic endowment effect* is defined in (8). We estimate Z using the state effects estimated in regression equation (6).

We compute the counterfactual average outcome of the Enslaved had their ancestors been distributed as the Free within the South. We find that the differential exposure to state-specific factors explains the vast majority of the Free-Enslaved gap even under this lower bound (see Appendix Table A.7). For example, had descendants of the Enslaved been distributed across Southern (but not Northern) states similar to the Free population, the Free-Enslaved gap would have been at least 67 percent smaller in 1940. Thus, the *geographic endowment effect* caused most of the persistence in the Free-Enslaved gap in the long run.

6.4 Location of Freedom and the Question of Exogeneity

Estimating the causal effect of place-specific factors requires that a person's location is orthogonal to their potential outcomes. Past research typically relies on "mover designs" (e.g., Chetty et al., 2016). In those studies, places' effects are estimated from the outcomes of families who move between them. Assumptions on the nature of their moves allow for a causal interpretation.

Our empirical strategy relies on a specific population's *immobility* rather than mobility. In particular, we build on the circumstance that the Enslaved did not enjoy the freedom of movement before 1865, leaving no room for self-selection into location. This circumstance lends plausibility to the key identifying assumption of an enslaved person's birthplace to be orthogonal to the potential outcomes of their (third-generation) descendants. The main threat to our identification assumption is the possibility of selective forced migration of enslaved people. Even though the Enslaved did not choose where they lived, owners' or traders' decisions may have induced selection into enslavement locations.

Slaveholder migration and the domestic slave trade contributed equally to the forced migration before 1865 (Fogel and Engerman, 1974; Tadman, 1979; Pritchett, 2001; Steckel and Ziebarth, 2013). Slaveholders were generally non-selective in moving all their enslaved people with them (Fogel and Engerman, 1974; Pritchett, 2001; Tadman, 2008; Pritchett, 2019). 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. The domestic slave trade accounts for the remaining inter-regional slave mobility. Selective slave trade is only evident in the small sugar cultivation areas.²⁵ Sugar cultivation accounted for 6 percent of the rural enslaved population (Tadman, 1977, 1979).²⁶

Overall, selective migration among the Enslaved may attenuate our estimates of place effects. If anything, one can hypothesize that the selection into location based on physical traits has biased upward the estimates of states that supposedly selected positively on height and strength. In contrast, we find that such states—those in the Lower South in general and those in the sugar region of Louisiana in particular—were especially detrimental to Black economic progress.

The results from the following section strongly support our key identifying assumption. Because our border discontinuity estimates capture the vast majority of differences in the causal state effects, any relevant selection would need to occur sharply at the border. Such forms of selection are implausible given that enslaved people were—if anything—selectively forced to migrate to specific locations based on the crops cultivated there. We verify that crops do not discontinuously change across state borders. We also verify that the observable characteristics of enslaved people—such as their age in 1860 or their literacy in 1870—did not discontinuously vary across borders, ruling out selection on observable characteristics directly.

7. THE JIM CROW EFFECT

Our analysis so far attributes the Free-Enslaved gap’s persistence primarily to the two groups’ differential exposure to state-level factors. This section assesses the role of Jim Crow regimes within these state factors. First, we use a border discontinuity design to show that geographic disparities in Black economic progress materialize sharply at state borders, implicating institutional factors, including Jim Crow, as main drivers. Second, we show that states’ Jim Crow regimes are a leading explanation underlying the impact of state institutions. Specifically, we find that 1) the negative impact of these state institu-

²⁵In contrast to the sugar industry, the cotton and tobacco industries (accounting for around 87 percent of enslaved agricultural workers) were generally non-selective on age and sex (Tadman, 1977).

²⁶By the nature of the work required, enslaved people there tended to be physically stronger and more likely to be male (Phillips, 1918). Traded enslaved people 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). Pritchett (2001) finds that traded enslaved people were marginally taller than the average enslaved population, conditional on age and sex, but Steckel and Ziebarth (2016) contest this finding. Physical characteristics were also co-determined by environmental influences such as nutrition, illness, or stress (Steckel, 1979; Carson, 2008). There is no evidence that traders selected enslaved people on anything other than such basic physical characteristics. This is consistent with the dehumanization of Black people that characterized the slave trade, which “reduced people to the sum of their biological parts” (Smallwood, 2008, p. 43).

tions was race-specific, largely leaving the economic status of white families unaffected, 2) the border discontinuities can be statistically explained by various measures of states' Jim Crow intensity, and 3) the timing of the effects from state institutions coincides with the rise of Jim Crow regimes.

7.1 State Institutions and Black Economic Progress After Slavery

Places may affect families' human capital (and other outcomes) for many reasons, be it cultural, climatic, economic, or institutional. We argue that only institutions change sharply at state borders, while other factors vary continuously. Therefore, to distinguish the effects of institutions from those of other factors, we decompose the location-specific parameters in equation (1):

$$\gamma_\ell^t = \gamma_{\epsilon(\ell)}^t + \gamma_{s(\ell)}^t, \quad (11)$$

where $\gamma_{\epsilon(\ell)}^t$ captures factors that vary continuously across state borders and $\gamma_{s(\ell)}^t$ captures factors that vary discontinuously across state borders. We can think of $\epsilon(\ell)$ as the geographic coordinates of location ℓ , and $s(\ell)$ as the state that location ℓ is in.²⁷ In the next section, we propose a border discontinuity design to separate the effect of institutions, $\gamma_{s(\ell)}^t$, from the effect of non-institutional factors, $\gamma_{\epsilon(\ell)}^t$.

7.2 Border Discontinuity Design

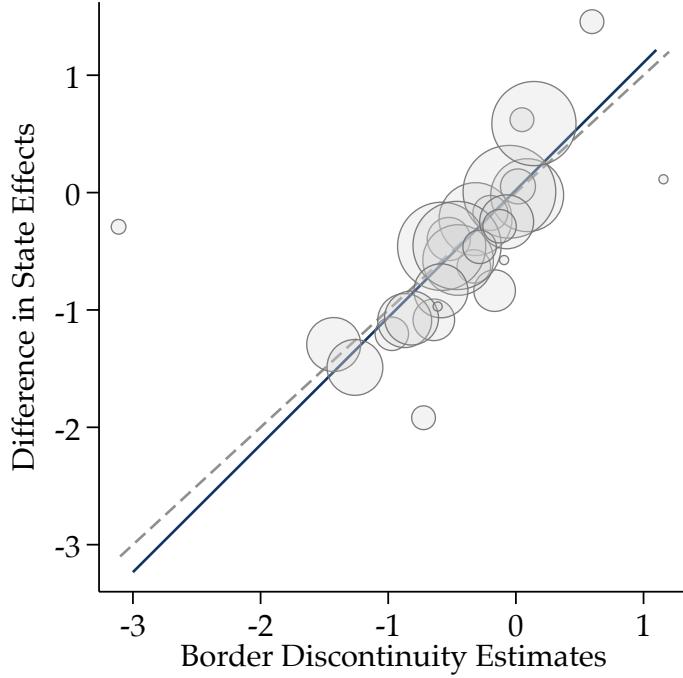
Our border discontinuity design compares the socioeconomic status of families in 1940 whose ancestors were freed on different sides of (but in close proximity to) state borders within the South in 1865 (see Appendix Figure B.21). We provide quantitative evidence that culture, climate, and economic activity do not vary discontinuously across state borders. We therefore argue that our border discontinuity design isolates the effect of institutions. The border discontinuity design takes the following form:

$$y_{i,b}^{1940} = \alpha_b + \beta_b \cdot \text{High}_{i,b}^{1870} + v_b \cdot \text{dist}_{i,b}^{1870} + \psi_b \cdot \text{dist}_{i,b}^{1870} \cdot \text{High}_{i,b}^{1870} + \varepsilon_{i,b}, \quad (12)$$

separately for each border b in the South, where $y_{i,b}^{1940}$ is the years of education of Black person i in 1940 whose ancestors were freed close to state-border b , $\text{High}_{i,b}^{1870}$ indicates whether i 's 1870 ancestors lived on the side of border b that had a more intensive Jim Crow regime than the state on the other side of the border, and $\text{dist}_{i,b}^{1870}$ is the distance between border b and the county's centroid in which i 's ancestors lived in 1870. The main coefficient of interest, β_b , captures the long-run effect of being freed on the more oppressive side of border b on a Black family's human capital.

²⁷Formally, $||\epsilon(\ell) - \epsilon(\ell')|| \rightarrow 0 \Rightarrow |\gamma_{\epsilon(\ell)}^t - \gamma_{\epsilon(\ell')}^t| \rightarrow 0$, whereas $\gamma_{s(\ell)}^t$ only depends on which side of a border ℓ is on, not on the precise coordinates $\epsilon(\ell)$: $\gamma_{s(\ell)}^t = \gamma_s^t$.

FIGURE 6: Differences in Black Economic Progress Arise Sharply at State Borders



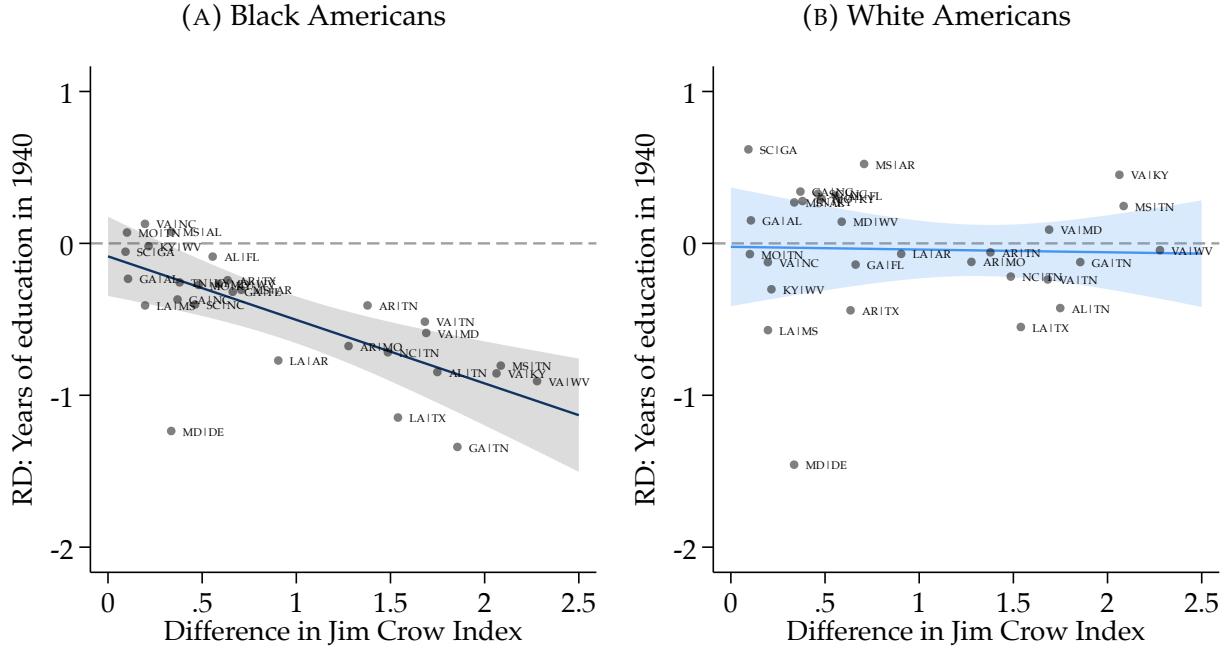
Notes: This figure relates each RD estimate (as shown in Figure 7) to the difference in the two states' overall causal effect on 1940 years of education (as shown in Panel A of Appendix Figure A.6). Estimates are weighted by the minimum sample size underlying the difference in state effects. A gray dashed 45 degree line shows the benchmark of equal differences across two states and across the border counties of two states. The blue line shows the best weighted linear fit ($\hat{\beta} = 1.08^{***}$, $R^2 = 0.75$). See Data Appendix C for details on the sample and data.

To assess to which degree institutions shaped the geography of Black economic progress, we compare the differences in progress that arise sharply at the border with the overall differences between states' effects (see Figure 6). We find that the border discontinuities are almost identical to differences in neighboring states' overall long-run effects on Black economic progress. This finding suggests that the geography of Black economic progress is mostly driven by institutional factors, less by factors that continuously vary across borders—such as economic activity, culture, or climate.

Having established the importance of state institutions, we next examine whether it was Jim Crow institutions specifically that slowed Black economic progress. To do so, we correlate our border discontinuity estimates $\hat{\beta}_b$ with differences in Jim Crow intensity, using that Jim Crow regimes differ more drastically across some borders than others. To quantify Jim Crow severity—which encompasses both de jure and de facto tactics (Woodward, 1955; Acemoglu and Robinson, 2008)—we employ a range of proxies that, despite their differing natures, are highly correlated. For example, the HRR index and the number of Jim Crow laws have a correlation of $\rho = 0.71$; the HRR index and Black school quality have a correlation of $\rho = -0.94$ (see Appendix Figure C.30). Across these measures, we consistently arrive at the same key finding.

We find that states' intensity of Jim Crow regimes predicts border discontinuities in

FIGURE 7: Regression Discontinuity Estimates and Jim Crow



Notes: Panel (A) of this figure shows each separate RD estimate in 1940 years of education for Black families whose ancestors were freed on different sides of state borders in 1865. Panel (B) shows the same for white families depending on where their ancestors lived in 1870. Each label shows the more oppressive before the less oppressive state. Negative estimates reflect lower education in more oppressive states. Lines show the best linear fit between RD estimates and the differences in Jim Crow intensity, weighted by the inverse of the estimates' standard error. Shaded areas represent robust 95 percent confidence bands. For point estimates, we use empirical Bayesian shrinkage as described in Appendix B.10. See Data Appendix C for details on the sample and data.

Black economic progress. Specifically, families freed in states with stricter regimes experienced significantly lower rates of progress, starting from the Jim Crow era (see Panel A of Figure 7). These gaps widen as the difference in Jim Crow severity increases across a border. For example, consistent with Louisiana's stricter Jim Crow regime compared to Texas's, families freed in Louisiana attained 1.2 fewer years of education by 1940 than those freed just miles away in Texas. Similarly, residing in states with more oppressive Jim Crow regimes led to lower incomes and a greater likelihood of working as a farmer in 1940 (see Appendix Figure A.14). No differences emerge for families freed across borders where states have comparable institutions. Incorporating extensive controls for 1860 local demographics, characteristics of slaves, crop suitability, and economic activity further accentuates these findings (see Appendix Figure B.19).

We also find that, as expected, families who left their enslavement state before the Jim Crow era were unaffected by their origin state's Jim Crow regime (see Appendix Figure A.16). However, if a family stayed and became exposed to the Jim Crow regime, the exposure had a persistent effect even for families who migrated in later decades. For instance, families freed in states with strict Jim Crow regimes and who stayed there until 1920 were still strongly impacted by their pre-1920 experiences in 1940. The longer a

family was exposed, the larger the effect on their socioeconomic status.

In principle, Jim Crow may have not only harmed Black Americans, but it could also have affected white Americans. First, some Jim Crow laws may have directly harmed poor white Americans. For example, poll taxes aimed at disenfranchising Black voters may also have disenfranchised some poor white voters. Second, Jim Crow may have benefited white elites. For example, vagrancy and emigrant-agent laws depressed farm workers' wages, potentially increasing land-owning families' profits.

We find that in contrast to Black families, the socioeconomic status of white families was not negatively affected by the Jim Crow intensity of the state in which their ancestors lived in 1870 (see Panel B of Figure 7). The same is true even for poor white Americans whose ancestors had no measurable human or physical capital in 1870 (see Panel (A) of Appendix Figure A.15). Our findings are consistent with existing evidence of Black Americans being the main beneficiaries of ending Jim Crow through the Civil Rights legislation (Wright, 2013).

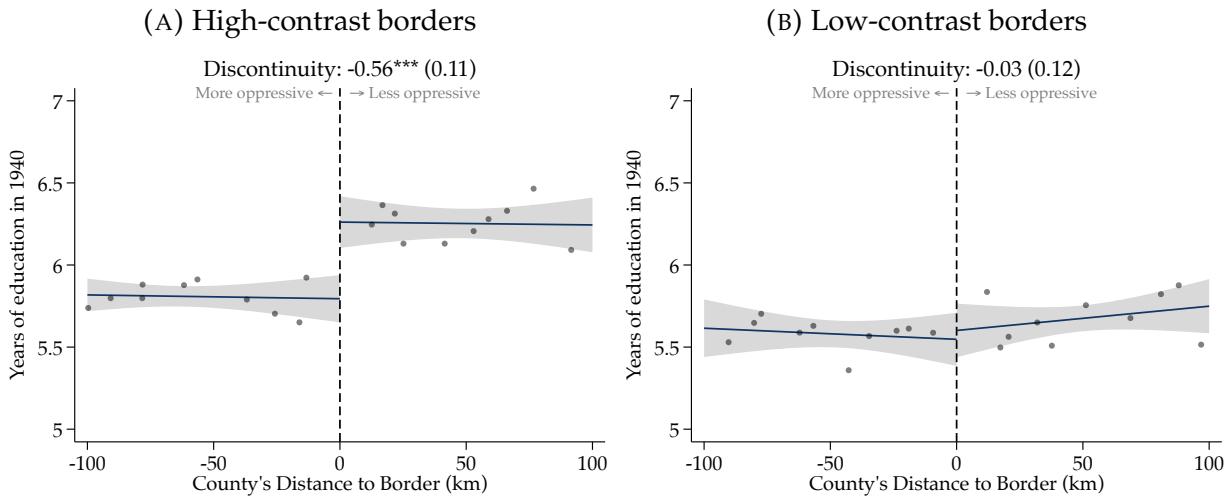
We do, however, find *positive* effects for the white land-owning elite. We find that the more oppressive a Jim Crow regime is, the more economically significant the gains by the border region's wealthiest ten percent of white families (see Panel (B) of Appendix Figure A.15). In sum, our results suggest that Jim Crow was an extractive institution that benefited the wealthiest white families at the cost of Black families while shielding poor white families from economic harm.

The end of slavery led to a drastic change in the geography of racially oppressive institutions in the US. State governments took the leading role in instituting Jim Crow regimes to limit the economic progress of newly freed enslaved families. Our results show that state institutions became a crucial determinant of how likely a Black family was to experience severe forms of oppression over the next century, shaping Black families' long-run economic progress. In the next section, we confirm that our border discontinuity design isolates the Jim Crow effect without being confounded by other factors.

7.3 Validation of the Border Discontinuity Design

To validate our border discontinuity design, we pool all borders, rather than estimating discontinuities for each border separately. The pooled regression equation closely follows equation (12). We equally divide our sample into two types of borders: "high-contrast borders" between states that strongly differ in their Jim Crow intensity (more than the median border difference in the HRR index); and "low-contrast borders" between states that differ less in their Jim Crow intensity (less than the median border difference). Our validation exercises focus on high-contrast borders, but the results generalize to the low-contrast borders.

FIGURE 8: Pooled Regression Discontinuity Estimates



Notes: This figure shows the RD estimates in 1940 years of education for Black families freed across state borders with different Jim Crow intensity in 1865. Panel (A) shows “high-contrast borders” where Jim Crow intensity differs more than across the median border (0.56 Jim Crow index points or more); Panel (B) shows “low-contrast borders” where it differs less than the median. The left half of each panel represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of the border population. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands clustered at the 1870 county level. See Data Appendix C for details on the sample and data.

Consistent with our main estimates, sharp educational differences only arise for Black families freed across borders where institutions differ substantially (see Figure 8).²⁸ Being freed on the more oppressive side of such a high-contrast border sharply reduced the years of education in 1940 by 0.6 years—10 percent of the average education among Black men.

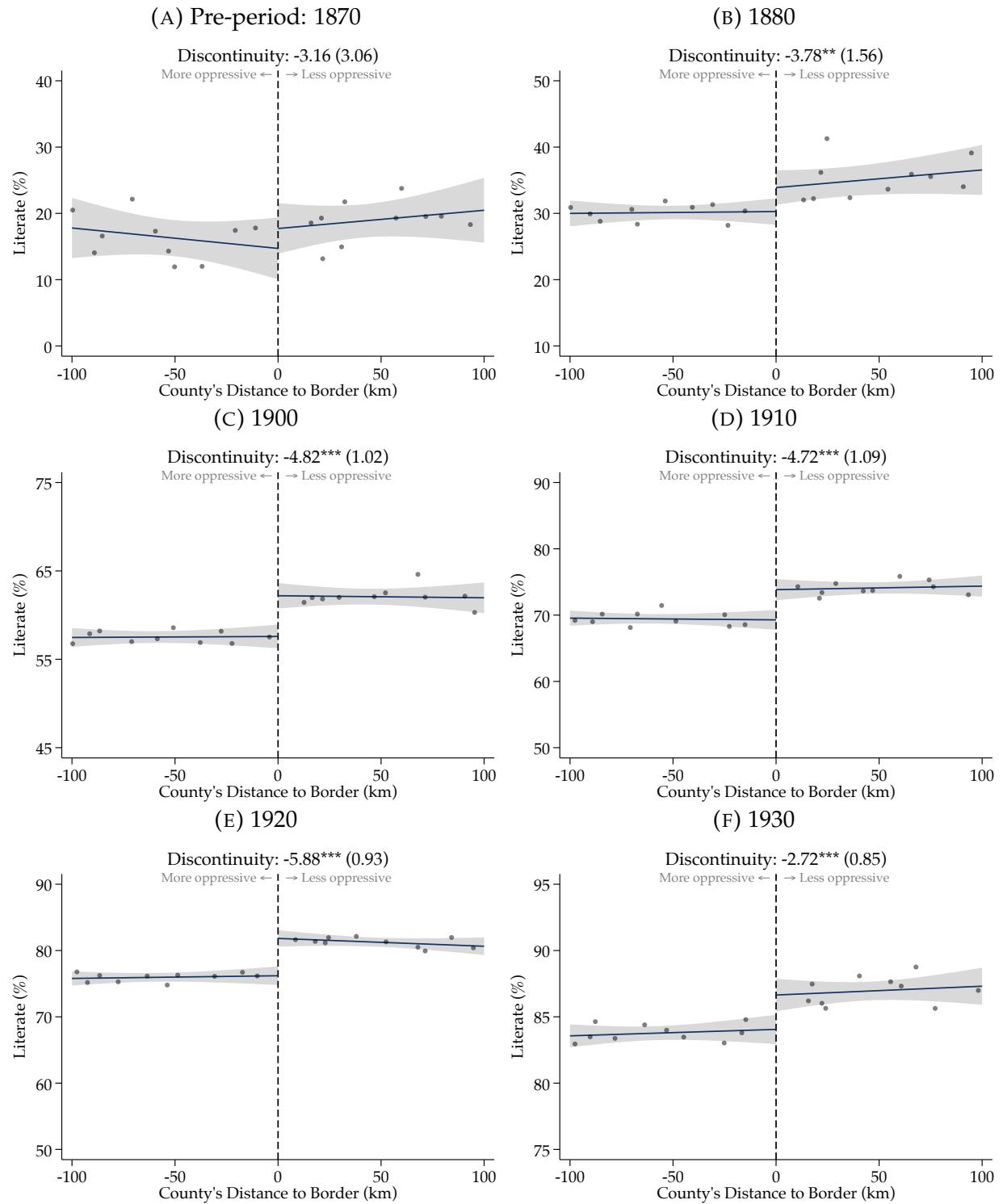
First, we confirm that differences across high-contrast borders only arise after the onset of Jim Crow (see Figure 9). Before Jim Crow, there were no differences in literacy among families freed in states that would become more oppressive during Jim Crow.²⁹ In 1880, the literacy rates of families were still equal. Starting in 1900, Black families attained lower literacy rates in more oppressive states. These differences grow over time in absolute terms but even more so in relative terms. By 1930, while almost 90 percent of all Southern Black men were literate, families freed in more oppressive states were still 4.6 percentage points less likely to be able to read and write.

Second, we confirm that before Jim Crow, location characteristics evolved smoothly across state borders. In 1860, neither the number of enslaved people relative to a county’s overall population, the share of its Black population, its cotton output per capita, nor its average farm value differed across state borders in the South (see Appendix Figure A.9). The same holds for other location characteristics such as population density, incomes, the age of enslaved people, and migration costs.

²⁸ Appendix Figure A.11 shows the pooled RD estimate for all borders—both high- and low-contrast.

²⁹ Appendix Figure A.8 shows RD estimates in literacy rates over time, separately by border.

FIGURE 9: Regression Discontinuities in Literacy (High-Contrast Borders)



Notes: This figure shows the RD estimate in literacy for Black families freed across state borders with different Jim Crow intensity in 1865. The sample is restricted to high-contrast borders (0.56 Jim Crow index points or more). The left half of each panel represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of the border population. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands clustered at the 1870 county level. See Data Appendix C for details on the sample and data.

Third, we address potential concerns that the impact of Jim Crow institutions may not vary significantly across states but rather within more localized geographies by presenting evidence to the contrary. Key outcomes directly influenced by Jim Crow policies exhibit marked gaps across state borders with differing Jim Crow intensities (see Appendix Figures C.34, A.12, and A.13). Specifically, counties in states with more severe Jim Crow regimes have sharply lower voter participation, Black school attendance, and Black teacher wages, plausibly reflecting the direct impact of suffrage restrictions and reduced school funding instituted in those states.

Fourth, our results are robust to using alternative measures for the intensity of states' Jim Crow regimes. We consider both the HRR index and a state's number of Jim Crow laws (see Appendix Figure B.20).

Last, we show that our results are robust to different cutoffs for the distance between a county's centroid and a state border: 100, 150, 200, or 250 kilometers (see Appendix Figure B.22). For example, the pooled RD estimates across high-contrast borders (as shown in panel (A) of Figure 8) for those cutoffs all range between -0.55 and -0.45 and are all highly significant. Our baseline bandwidth is 100 kilometers—close to the mean squared error (MSE)-optimum.

The results from our regression discontinuity design also strongly support our key identifying assumption—that the birthplace of an enslaved person is orthogonal to their innate “ability.” Specifically, we find that the differences in the causal effects of states sharply and fully arise at state borders. Therefore, the main potential threat of selection bias remains the selection of enslaved people into states sharply around borders. However, any plausible selection into the destination of forced migration was based on the crop cultivated in an area that, as we confirm, transcends state borders (along with many other characteristics of border areas). Therefore, the selection of enslaved people into location is implausible to affect our results. In addition, we directly rule out selection based on observable characteristics, showing that the characteristics of enslaved people, such as their age during or their literacy immediately after slavery, do not differ across borders. We also find that the number of lynchings between 1883 and 1941 does not vary sharply across borders, supporting the assumption that border differences in economic progress capture the effect of state institutions (see Appendix Figure A.17).

In sum, our evidence suggests that states' Jim Crow regimes played a critical role in shaping the South's detrimental effect on Black economic progress. The estimates are a lower bound for Jim Crow's importance because all Southern states adopted Jim Crow regimes. Our estimates only isolate the *additional* effect of more oppressive institutions rather than their aggregate effects.

8. THE MECHANISM OF LIMITED ACCESS TO EDUCATION

Leading scholars have pointed out the importance of Jim Crow in limiting Black families' long-run human capital accumulation. Booker T. Washington writes that "few people [have an] idea of the intensive desire which [Black people] showed for education. It was a whole race trying to go to school" ([Washington, 1907](#)). However, Black people's desire for education was met with resistance. "[Black Americans'] attempts at education provoked the most intense and bitter hostilities as evincing a desire to render themselves equal to the whites" (Freedmen's Commission Report cited in [Du Bois, 1935](#), p. 645). Robert Higgs argues that governments were the leading force of this resistance:

"Most damaging of all [racial discrimination after slavery] was the discriminatory behavior of the southern state and local governments. By providing only scant resources for black education, public school boards helped to perpetuate illiteracy [...], and they thereby set in motion a variety of adverse effects." ([Higgs, 1989](#), p. 25)

We first use our newly built database on laws and their content to explore the relative importance of different domains that Jim Crow regimes affected. We document that the most significant number of laws pertained to education, accounting for one-third of all Jim Crow laws passed across the South until 1950 (see Appendix Figure [C.31](#)).³⁰

Jim Crow laws on education established the provision of resources for new schools or colleges for white Americans only. They also required the racial segregation of existing schools or local school boards to comprise only white people. Even school books were regulated, stipulating that once a Black or white child had used a book, children of the other race were not allowed to use the same book. Those laws likely created drastic differences in the educational resources available to Black and white children. Indeed, we find a robust negative correlation between a state's number of education-specific Jim Crow laws and the quality of Black schools ($\rho = -0.70$).

Our analysis of Black teacher wages confirms that disparities in school quality are pronounced right at states' borders, underlining the critical role of institutional factors in shaping the quality of Black schools (see Appendix Figure [C.34](#) and [Naidu, 2012](#); [Card et al., 2022](#)). We also explore the importance of education-specific Jim Crow regimes for Black economic progress by repeating our regression discontinuity design based on the number of education-specific Jim Crow laws and the quality of Black schools ([Card](#)

³⁰A category's number of Jim Crow laws is not a conclusive measure of its importance; suffrage laws are a prime example. Suffrage laws are low in number, but their effects are massive (see e.g., [Naidu, 2012](#)). Laws in other categories are likely a downstream outcome of Black voter disenfranchisement ([Engerman and Sokoloff, 2011](#)). Therefore, while the number of Jim Crow laws on education is extensive, only through the following analysis can we conclude that they were a crucial part of states' Jim Crow regimes.

and Krueger, 1992; Carruthers and Wanamaker, 2017). Both measures capture the sharp differences in Black economic progress across Jim Crow regimes (see Appendix Figure A.10). The finding is consistent with Card and Krueger (1996) and Card et al. (2022) who show that state-level factors induced critical differences in school quality, especially for Black children, “helping to explain the persistence of the human capital gap between Blacks and whites.”

9. CONCLUSION

This paper provides new evidence on the long-run impact of racially oppressive institutions, finding that Black Americans’ economic status today depends strongly on their historical exposure to those institutions. First, we document that Black families enslaved until 1865 continue to have considerably lower education, income, and wealth today. Second, we show that this persistence is driven by post-slavery oppression under Jim Crow and discuss Black Americans’ limited access to education as a critical mechanism.

We have put forward a new framework for slavery’s legacy to incorporate systemic discrimination of the formerly Enslaved and their descendants under Jim Crow. The institution of slavery determined *where* a Black family was freed from slavery. We show that the state where a family was freed determined the Jim Crow regime they likely faced over the subsequent decades. While Jim Crow compressed the socioeconomic status of Black Americans *within* states, differences in Jim Crow intensity led to pronounced disparities *across* states, thereby placing descendants of those enslaved until the Civil War at a disproportionate disadvantage. After 1940, the main reason descendants of families enslaved until the Civil War have lower socioeconomic status is their concentration in the states that adopted the most strict Jim Crow regimes starting in 1877. Systemic discrimination—the higher exposure to ongoing discrimination *because of past discrimination* (Bohren et al., 2022)—is a central aspect of slavery’s persisting legacy.

Despite the end of Jim Crow, today’s geography of Black economic progress has similarities with that of the past. States that impeded Black economic progress post-slavery also limit intergenerational mobility for low-income children today (see Appendix Figure A.7 and Berger, 2018). However, different from the Jim Crow era, those differences do not arise sharply across state borders. Future research should investigate why places’ capacity to generate upward mobility has persisted despite drastic institutional change. Part of the answer may lie in anti-Black resentment, which remains high in places with historical prevalence of slavery and Jim Crow (Acharya et al., 2018).

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 *within-race* disparities that race-specific policies may not address. College affirmative

action is a prime example. [Massey et al. \(2007\)](#) show that the more selective a college, the less likely Black students are to descend from the Enslaved. While only 13 percent of 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, 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 Black families today are impacted drastically 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 Lower South compared to those who gained freedom earlier, mainly in the Upper South and North. Many free Black Americans had been enslaved in earlier periods, and all Black Americans faced discrimination due to slavery and Jim Crow 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.

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 unavailable or have poor coverage for women. Women have historically tended to change their surnames upon marriage, making it impossible for conventional methods to link them across census records. While not within the scope of this paper, we link millions of women across census records in a separate project by using the information on their maiden and married names from social security applications ([Althoff et al., 2023](#)). Second, we emphasize the significance of educational Jim Crow institutions as a crucial mechanism; however, institutions related to other aspects may have further impeded Black economic advancement. Although several of these institutions have been thoroughly investigated (e.g., restrictions on Black suffrage—see [Naidu, 2012](#)), numerous others remain relatively unexplored (e.g., constraints on interracial marriage). This situation presents a wealth of promising opportunities for future research.

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, 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.
- ABRAMITZKY, R., J. CONWAY, R. MILL, AND L. STEIN (2023): "The Gendered Impacts of Perceived Skin Tone: Evidence from African-American Siblings in 1870–1940," Working paper.
- 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.
- ACEMOGLU, D. AND J. A. ROBINSON (2008): "Persistence of Power, Elites, and Institutions," *American Economic Review*, 98, 267–93.
- 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. (2021): "Two Steps Forward, One Step Back: Racial Income Gaps among Women since 1950," Working paper.
- ALTHOFF, L., H. BROOKES GRAY, AND H. REICHARDT (2023): "The Missing Link(s): Women and Intergenerational Mobility," Work in progress.
- ANGRIST, J. D. AND J.-S. PISCHKE (2008): *Mostly harmless econometrics*, Princeton university press.
- ASHENFELTER, O. AND A. KRUEGER (1994): "Estimates of the Economic Return to Schooling from a New Sample of Twins," *The American Economic Review*, 84, 1157–1173.

- 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.
- BLACKMON, D. A. (2009): *Slavery by another name: The re-enslavement of Black Americans from the Civil War to World War II*, Anchor.
- BOERMA, J. AND L. KARABAROUNIS (2021): "Reparations and Persistent Racial Wealth Gaps," Working Paper 28468, National Bureau of Economic Research.
- BOHREN, J. A., P. HULL, AND A. IMAS (2022): "Systemic Discrimination: Theory and Measurement," Working Paper 29820, National Bureau of Economic Research.
- BONILLA-SILVA, E. (2015): "The Structure of Racism in Color-Blind, "Post-Racial" America," *American Behavioral Scientist*, 59, 1358–1376.
- BOUSTAN, L. (2016): *Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets*, National Bureau of Economic Research Publications, Princeton University Press.
- CARD, D., C. DOMNISORU, AND L. TAYLOR (2022): "The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility," *Journal of Labor Economics*, 40, 39–95.
- CARD, D. AND A. B. KRUEGER (1992): "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics*, 107, 151–200.
- (1996): "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina," *The Journal of Economic Perspectives*, 10, 31–50.
- CARRINGTON, W. J., E. DETRAGIACHE, AND T. VISHWANATH (1996): "Migration with Endogenous Moving Costs," *American Economic Review*, 86, 909–930.

CARRUTHERS, C. K. AND M. H. WANAMAKER (2017): "Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap," *Journal of Labor Economics*, 35, 655–696.

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.

CASCIO, E. U. AND E. G. LEWIS (2022): "Legal Activism, State Policy, and Racial Inequality in Teacher Salaries and Educational Attainment in the Mid-Century American South," Working Paper 30631, National Bureau of Economic Research.

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, 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.

COHEN, W. (1991): *At Freedom's Edge: Black Mobility and the Southern White Quest for Racial Control, 1861–1915*.

COLLINS, W. J. (1997): "When the Tide Turned: Immigration and the Delay of the Great Black Migration," *The Journal of Economic History*, 57, 607–632.

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.

CUTLER, D. M., E. L. GLAESER, AND J. L. VIGDOR (1999): "The Rise and Decline of the American Ghetto," *Journal of Political Economy*, 107, 455–506.

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.

- 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*.
- ENGELMAN, S. L. AND K. L. SOKOLOFF (2011): *Economic Development in the Americas since 1500: Endowments and Institutions*, Cambridge University Press.
- 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.
- (2014): *Reconstruction: America's Unfinished Revolution, 1863-1877*, Harper Collins.
- FRAZIER, E. F. (1949): *The Negro in the United States.*, The Macmillan Company.
- 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/>.
- HIGGS, R. (1989): "Black progress and the persistence of racial economic inequalities, 1865-1940," in *The Question of Discrimination: Racial inequality in the US labor market*, Middletown, CT: Wesleyan University Press, 9–31.
- KOUSSER, J. M. (1974): *The Shaping of Southern Politics: Suffrage Restriction and the Establishment of the One-Party South, 1880-1910*, Yale University Press.
- LEEUWEN, M. H. D. V. AND I. MAAS (2011): *HISCLASS: A Historical International Social Class Scheme*, Universitaire Pers Leuven.

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

- SOWELL, T. (1978): "Essays and data on American ethnic groups," .
- STECKEL, R. H. (1979): "Slave height profiles from coastwise manifests," *Explorations in Economic History*, 16, 363.
- STECKEL, R. H. AND N. ZIEBARTH (2013): "A Troublesome Statistic: Traders and Coastal Shipments in the Westward Movement of Slaves," *The Journal of Economic History*, 73, 792–809.
- (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.
- U.S. SENATE (2019): "Landmark Legislation: The Civil Rights Act of 1964," .
- WALTON, H., S. C. PUCKETT, AND D. R. DESKINS (2012): *The African American Electorate: A Statistical History*, Washington, D.C.: CQ Press.
- WASHINGTON, B. T. (1907): *Up From Slavery: An Autobiography*.
- 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.
- WRIGHT, G. (2013): *Sharing the Prize: The Economics of the Civil Rights Revolution in the American South*, Harvard University Press.

APPENDIX

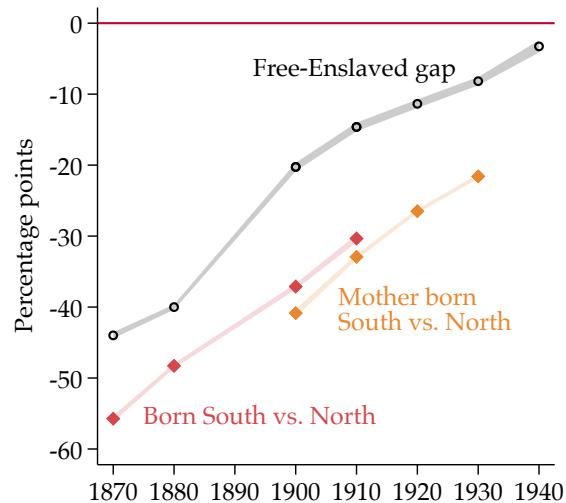
A Additional Results	47
A.1 Free-Enslaved Gap and a Literature Benchmark	47
A.2 Free-Enslaved Gap in Alternative Measures	48
A.3 Free-Enslaved Gap for Free Without Physical or Human Capital	50
A.4 Free-Enslaved Gap between and within Ancestor's Birthplace	51
A.5 State-Level Place Effects	54
A.6 Persistence of Place Effects	55
A.7 Border Discontinuities	56
B Robustness Checks	63
B.1 Free-Enslaved Gap Controlling for Skin-Tone	63
B.2 Free-Enslaved Gap Based on the Distribution of Surnames	64
B.3 Free-Enslaved Gap Excluding Places with Excess Mortality	65
B.4 Border Discontinuities Controlling for Location Characteristics	66
B.5 Border Discontinuities and Alternative Jim Crow Measures	67
B.6 Border Discontinuities Using Different Bandwidths	68
B.7 Adjusting Estimates for Intermarriage	69
B.8 Adjusting Estimates for Misclassification Bias	72
B.9 The Direct Effect of Locations After Accounting for Migration	75
B.10 Empirical Bayes Shrinkage	77
B.11 Placebo Exercises	80
B.12 Convergence Benchmark	82
C Data Appendix	84
C.1 Jim Crow Database	84
C.2 Census Data	90
C.3 Census Linking	90
C.4 Identifying Descendants of the Free and Enslaved	91

C.5 Main Sample	98
C.6 Credit Bureau Sample	100
C.7 Individual-Level Outcome Variables	100
C.8 Neighborhood-Level Outcome Variables	102
C.9 County Characteristics	104
C.10 Nine-Digit ZIP to Census 2000 Crosswalks	105
C.11 Descriptive Statistics	107
D Model Appendix	113

A. ADDITIONAL RESULTS

A.1 Free-Enslaved Gap and a Literature Benchmark

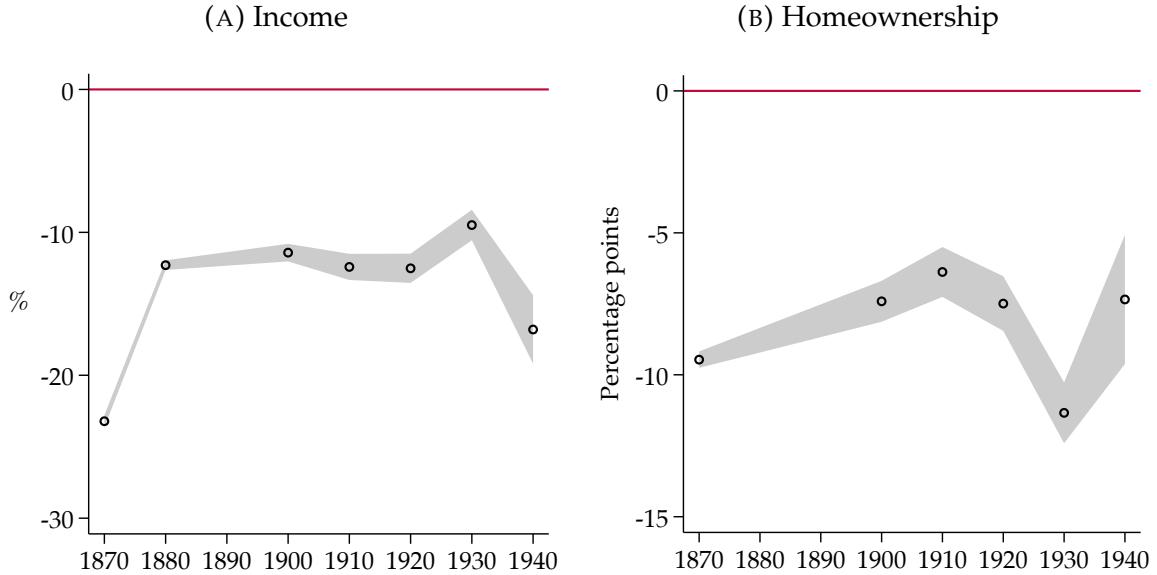
FIGURE A.1: Free-Enslaved and Southern-Northern Born Gap in Literacy (1870–1940)



Notes: This figure shows the gaps in literacy among prime-age (20-54) male descendants of free and enslaved Black Americans, as well as those born in the North and South, over each census decade. The gap between Southern and Northern-born individuals is estimated using full census data (not requiring record linkage) that include birthplaces or maternal birthplaces. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. For the Free-Enslaved gap, 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. See Data Appendix C for details on the sample and data.

A.2 Free-Enslaved Gap in Alternative Measures

FIGURE A.2: Free-Enslaved Gap (1870–1940)



Notes: This figure shows the gaps in income (occupational income score) and homeownership among prime-age (20-54) male descendants of enslaved vs. free Black Americans in each census decade. The sample includes both the South and North of the US. 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. See Data Appendix C for details on the sample and data.

TABLE A.1: Free-Enslaved Gap (1940) in Different Income Measures

	OCCSCORE (1950-\$) Mean: 1,604.09	LIDO Score (1950-\$) Mean: 1,161.69	Wage Income (1940-\$) Mean: 381.20	Total Income (1940-\$) Mean: 793.47	Song et al. Score Mean: 43.42
Ancestor Enslaved	-148.39*** (10.86)	-279.00*** (8.59)	-145.92*** (6.13)	-204.29*** (10.29)	-9.29*** (0.39)
Controls (age, age ²)	Y	Y	Y	Y	Y
Adjusted R ²	0.04	0.04	0.05	0.09	0.01
Observations	168,138	142,743	154,463	146,871	168,138
Ancestor Free	9,325	7,517	8,551	8,100	9,325

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, total predicted income, and the Song et al. score. We compute the Song et al. score by computing the average literacy rate by occupation and birth decade and converting this measure into ranks. The sample includes both the South and North of the US. All estimates are for Black prime-age men in 1940. Sample means are computed for the combined sample of the Free and Enslaved. See Data Appendix C 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 A.2: Free-Enslaved Gap (1940) in Different Education 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*** (0.26)	-1.59*** (0.05)	-7.86*** (0.45)	-1.86*** (0.21)	-0.74*** (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
<i>Ancestor Free</i>	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. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. The sample includes both the South and North of the US. All estimates are for Black prime-age men in 1940. Sample means are computed for the combined sample of the Free and Enslaved. See Data Appendix C 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 A.3: Free-Enslaved Gap (2000) using Mortality Records

	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*** (0.51)	-2.45*** (0.55)	-4,795.93*** (636.79)	-15,755.30*** (2,462.82)
Level of outcome	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
<i>Ancestor Free</i>	1,713	1,713	1,715	1,634

Notes: This table shows the Free-Enslaved gap in 2000 neighborhood-level outcomes: high school and college degrees, median incomes, and median house values (conditional on ownership). A neighborhood is a census tract. Each person is assigned the value of the census tract in which they last lived according to administrative mortality records. The sample includes both the South and North of the US. Sample means are computed for the combined sample of the Free and Enslaved. See Data Appendix C for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. ***
 $p < 0.01$, **
 $p < 0.05$, *
 $p < 0.1$.

A.3 Free-Enslaved Gap for Free Without Physical or Human Capital

TABLE A.4: Free-Enslaved Gap (1940) for Free Without Physical or Human Capital in 1860

	Education (Years)		Wage Income (USD)		Homeownership (%)		House Value (USD)	
	Mean: 5.83		Mean: 381.64		Mean: 29.08		Mean: 1,380.43	
Ancestor Enslaved	-1.00*** (0.15)	-0.12 (0.15)	-90.43*** (21.13)	26.85 (21.44)	-6.16*** (1.95)	-1.42 (2.00)	-343.74** (159.58)	440.28** (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
<i>Ancestor Free</i>	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 enslaved vs. free 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. The sample includes both the South and North of the US. 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 Enslaved. Appendix Figure A.3 shows the evolution of the conditional Free-Enslaved gap over time. See Data Appendix C for details. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A.4 Free-Enslaved Gap between and within Ancestor's Birthplace

FIGURE A.3: Free-Enslaved Gap Conditional on Ancestor State (1870-1940)

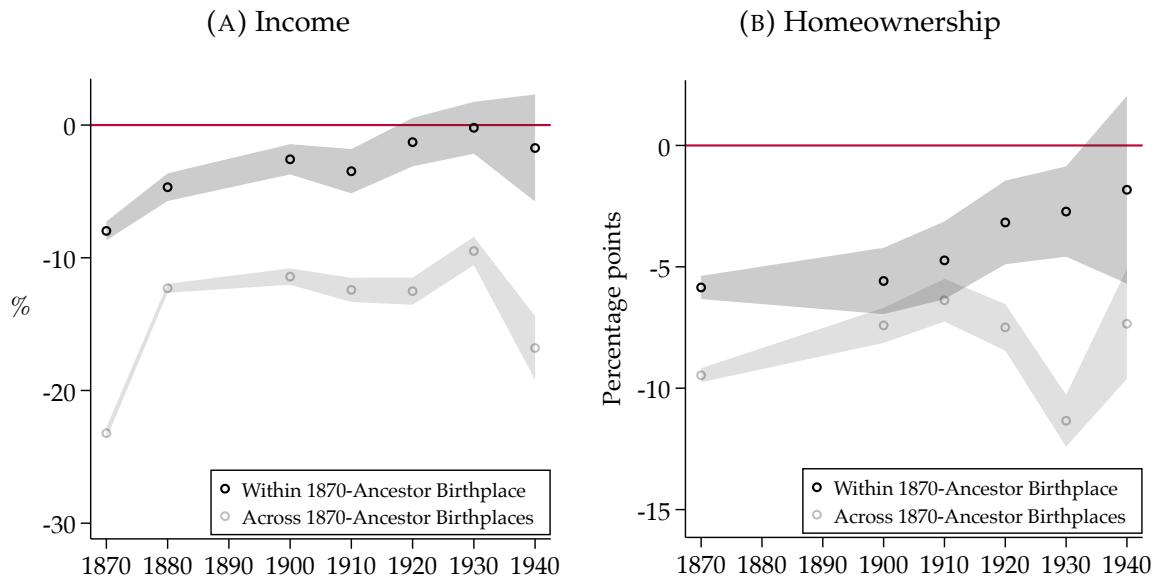
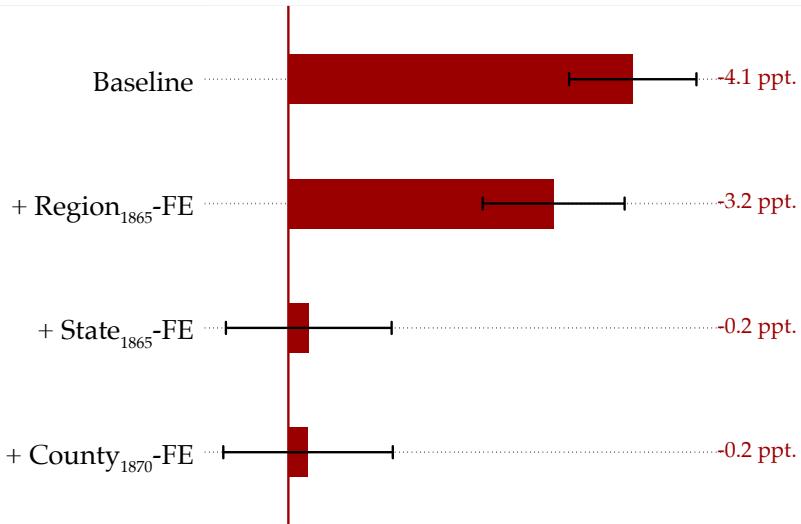


FIGURE A.4: Free-Enslaved Gap in Literacy Conditional on Ancestor Location (1940)



Notes: This figure shows the 1940 Free-Enslaved gap in literacy before and after including different levels of origin location fixed effects. We successively add fixed effects for the region (South or North) and state a family's 1870 ancestor were born, and the county in which their 1870 ancestors lived. The sample includes only Black prime-age (20–54) men whose ancestors can be located in 1870. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. See Data Appendix C for details on the sample and data.

TABLE A.5: Free-Enslaved Gap (1940) between and within Ancestor's Birthplace

	Education (Years)		Wage Income (USD)		Home Ownership (%)		House Value (USD)	
	Mean: 5.91		Mean: 388.01		Mean: 29.48		Mean: 1,412.17	
Ancestor Enslaved	-1.49*** (0.07)	-0.41*** (0.08)	-137.00*** (8.51)	-20.22** (9.84)	-6.76*** (0.86)	-1.61 (1.04)	-574.06*** (90.08)	8.40 (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
<i>Ancestor Free</i>	4,617	4,617	4,371	4,371	4,640	4,640	1,624	1,624

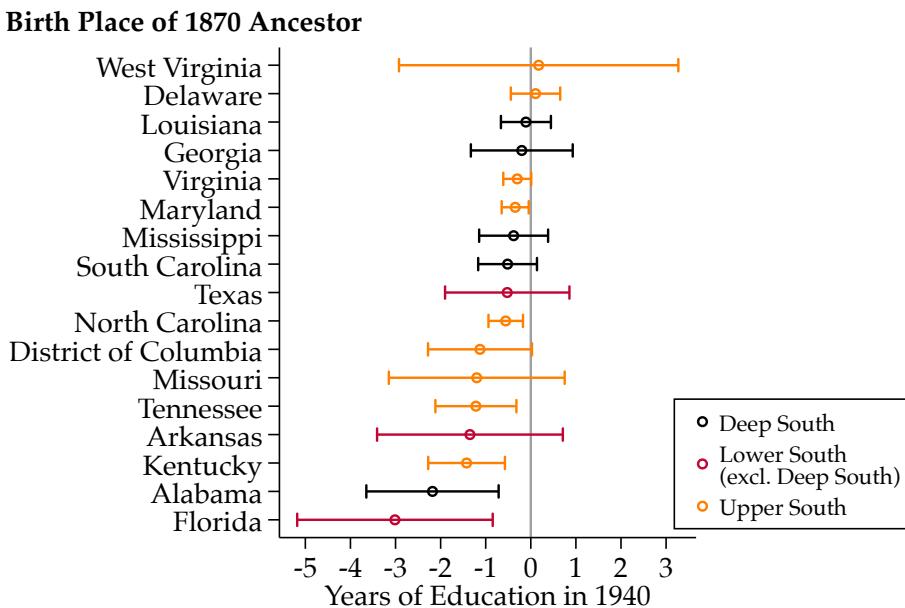
Notes: This table shows the gap in years of education, total income, homeownership rate, and house value among prime-age (20-54) male descendants of enslaved vs. free Black Americans in 1940. The sample includes both the South and North of the US. Only observations that can be linked to the 1850, 1860, 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 Enslaved. Figure 5 and Appendix Figure A.3 show the evolution of the conditional Free-Enslaved gap over time. See Data Appendix C 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 A.6: Free-Enslaved Gap (2000) at Census Block Level between and within Ancestor's Birthplace

	HS Degree (%)		College Degree (%)		Income (USD)		House Value (USD)	
	Mean: 69.20		Mean: 12.32		Mean: 30,143.90		Mean: 88,830.12	
Ancestor Enslaved	-2.57*** (0.74)	-0.89 (0.82)	-2.07*** (0.78)	-0.29 (0.78)	-5,032.50*** (921.89)	-1,014.92 (1,005.32)	-13,391.02*** (3,498.95)	-780.04 (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			N Y	N Y
Controls (age, age ²)	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
<i>Ancestor Free</i>	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. The sample includes both the South and North of the US. Columns 1, 3, 5, and 7 repeat Table A.3 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 Enslaved. See Data Appendix C 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 A.5: 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. The sample includes both the South and North of the US. 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.

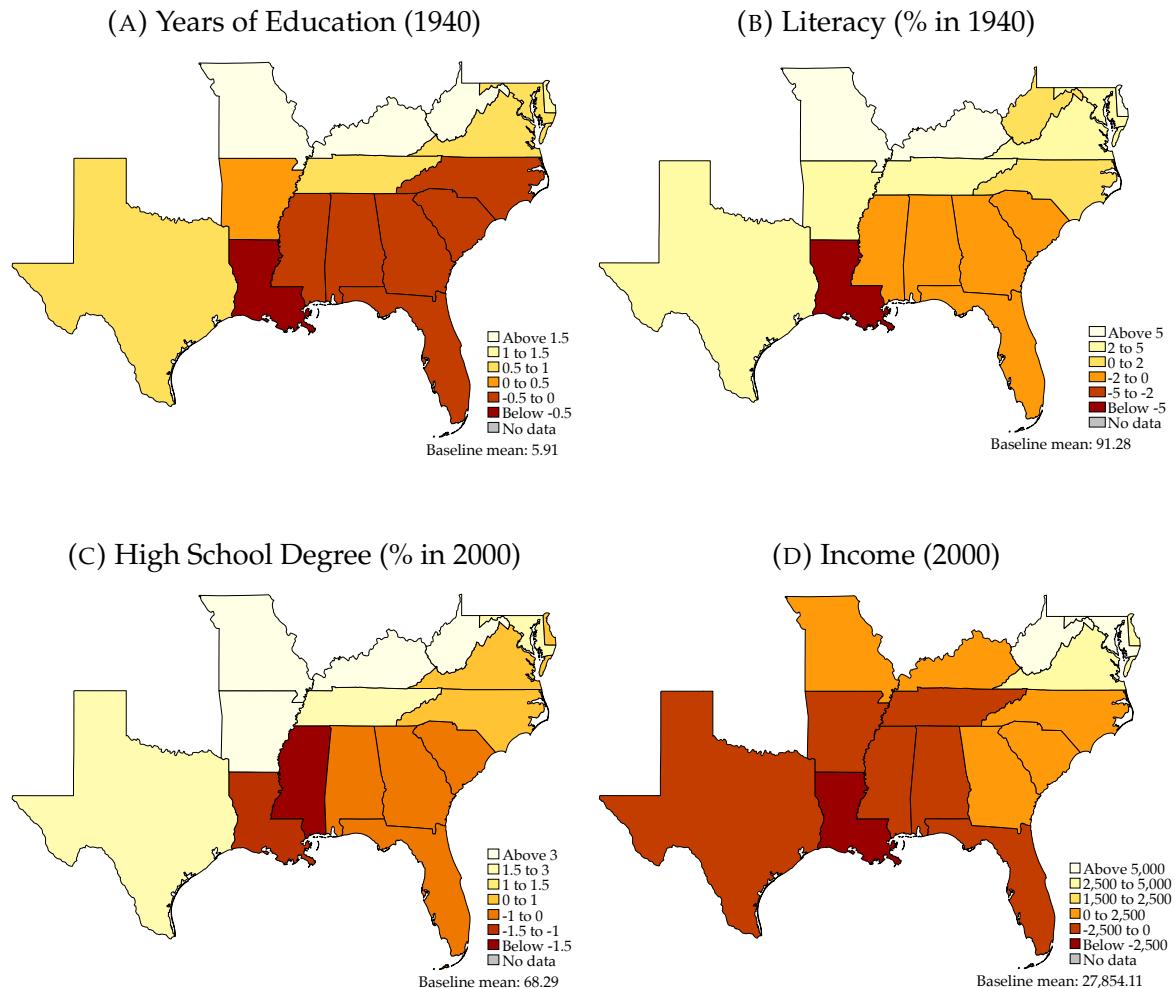
TABLE A.7: Decomposition of the Free-Enslaved Gap in 1940

	Free-Enslaved gap & ancestor location			Geography's effect as % of gap		
	National	Within South	Within state	Less conservative	Conservative	Lower bound
Literacy (%)	-4.2	-3.2	-0.4	138%	90%	67%
Years of education	-1.6	-1.2	-0.4	113%	75%	50%

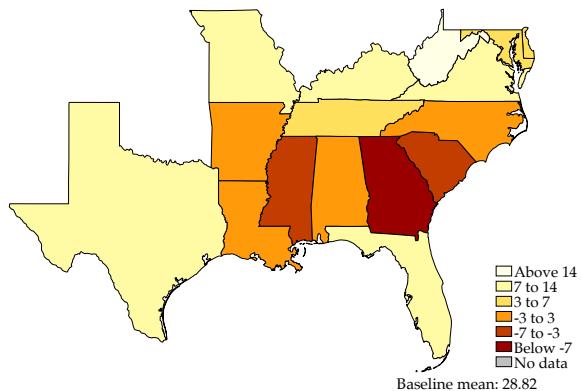
Notes: This table decomposes the 1940 Free-Enslaved gaps in literacy and years of education. We successively add fixed effects for the region (South or North) and state a family's 1870 ancestor were born, and the county in which their 1870 ancestors lived. Columns 4 and 5 show the fraction of the national Free-Enslaved gap (column 1) that can be accounted for by state variation (column 3), respectively including (less conservative) or excluding (conservative) extrapolated effects for the North. Column 6 shows the result of our formal decomposition in Section 6.3. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. The sample includes only Black prime-age (20–54) men whose ancestors can be located in 1870. See Data Appendix C for details on the sample and data.

A.5 State-Level Place Effects

FIGURE A.6: Long-Run Effect of Ancestor's State of Emancipation on Outcomes



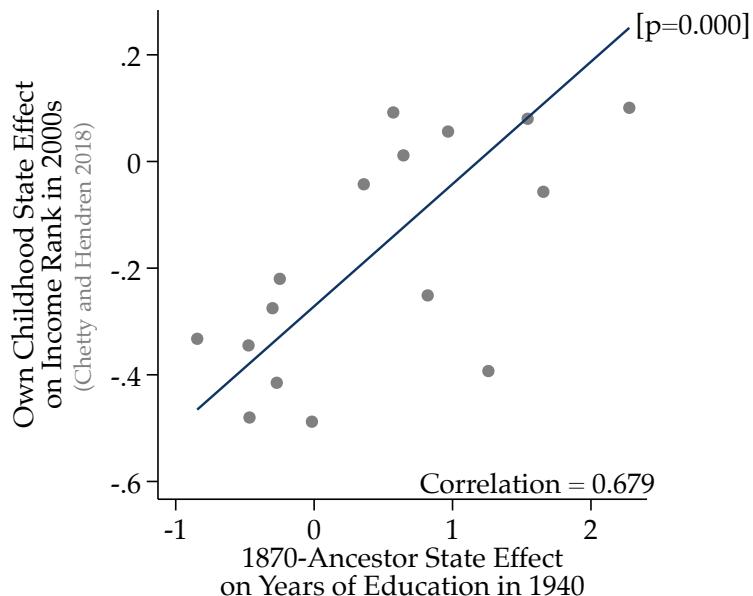
(E) Homeownership Rate (% in 1940)



Notes: This figure shows the 1870 ancestor state of birth fixed effect estimates on years of education and literacy rates in 1940, high school completion rates in 2000, homeownership rates in 1940, and income in 2000. 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. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. The sample includes Black prime-age (20–54) men whose ancestors can be located in 1870. See Data Appendix C for details on the sample and data.

A.6 Persistence of Place Effects

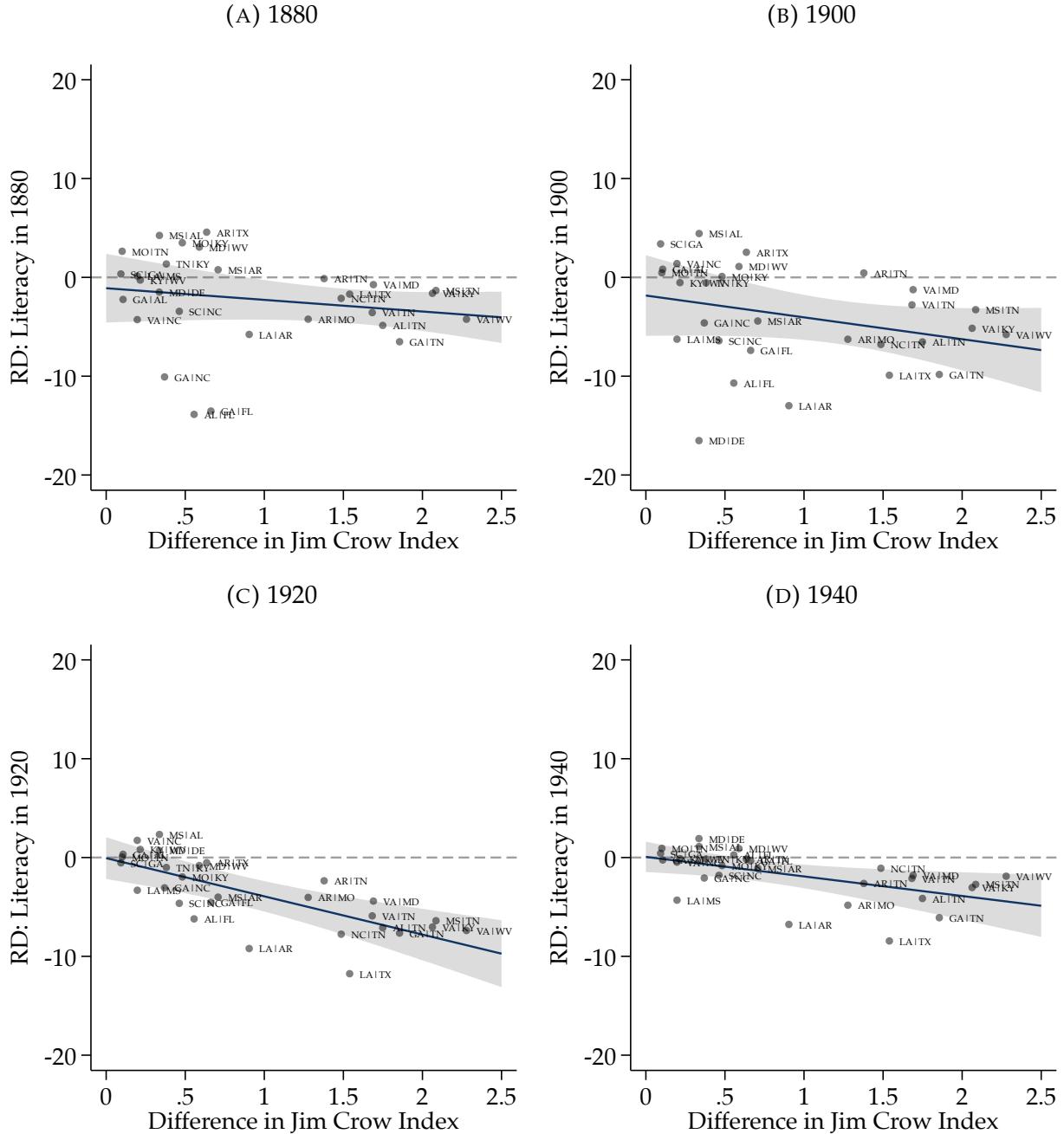
FIGURE A.7: Persistence of a State's Capacity to Generate Upward Mobility



Notes: This figure is a binned scatter plot relating a state's causal effect on Black economic progress from 1865 to 1940 (as shown in Panel A of Appendix Figure A.6) to the state's causal effect on intergenerational mobility in recent decades (as estimated by Chetty and Hendren, 2018). The modern estimates reflect a child's mean percentile rank in the national household income distribution at age 26 conditional on growing up with parents at the 25th percentile. See Data Appendix C for details on the sample and data.

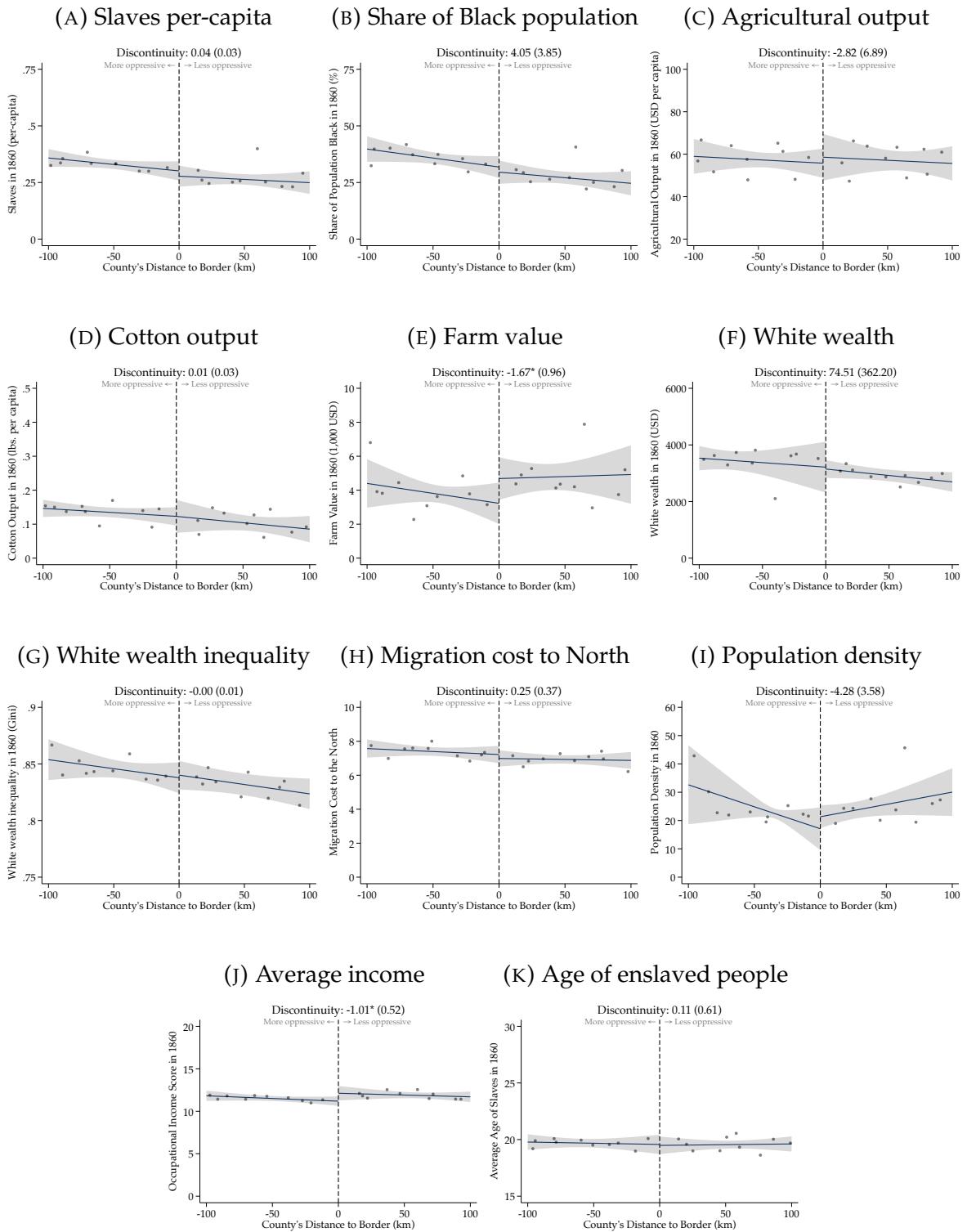
A.7 Border Discontinuities

FIGURE A.8: RD Estimates in Literacy over Time



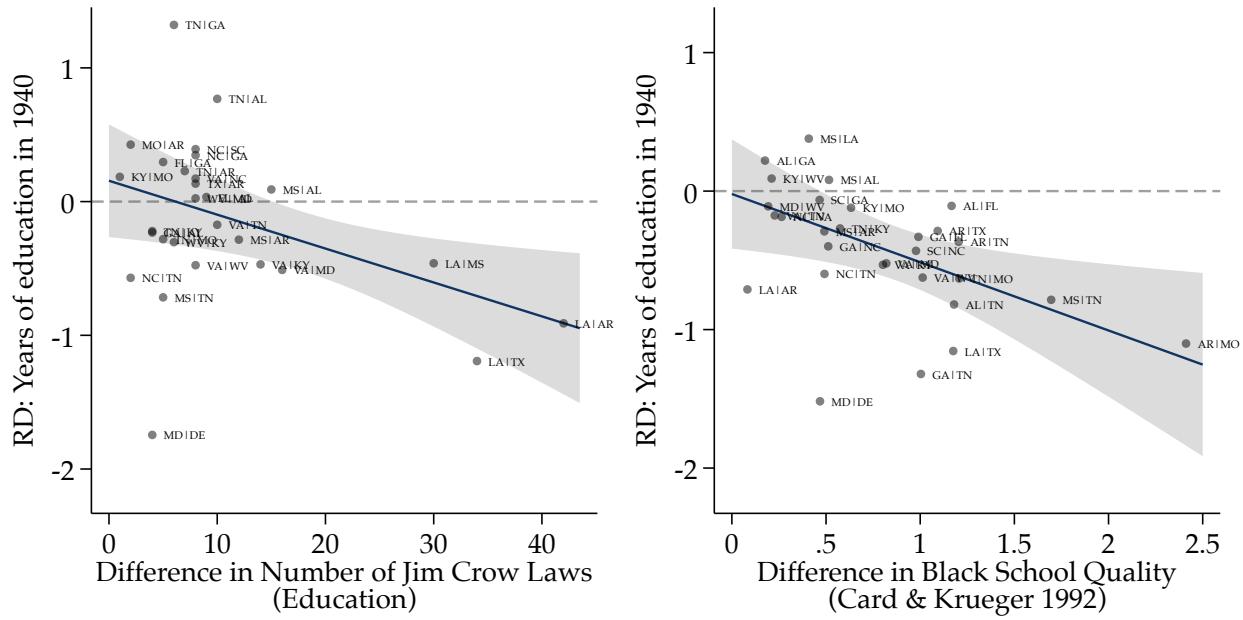
Notes: This Figure shows each separate RD estimate in literacy in 1880, 1900, 1920, and 1940 for Black families whose ancestors were freed on different sides of state borders in 1865. Each label shows the more oppressive before the less oppressive state. Negative estimates reflect lower literacy in the more oppressive state. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. Lines show the best linear fit between RD estimates and the differences in Jim Crow intensity, weighted by the inverse of the estimates' standard error. Shaded areas represent robust 95 percent confidence bands. For point estimates, we use empirical Bayesian shrinkage as described in Appendix B.10. See Data Appendix C for details on the sample and data.

FIGURE A.9: No Border Discontinuities in 1860 Location Characteristics



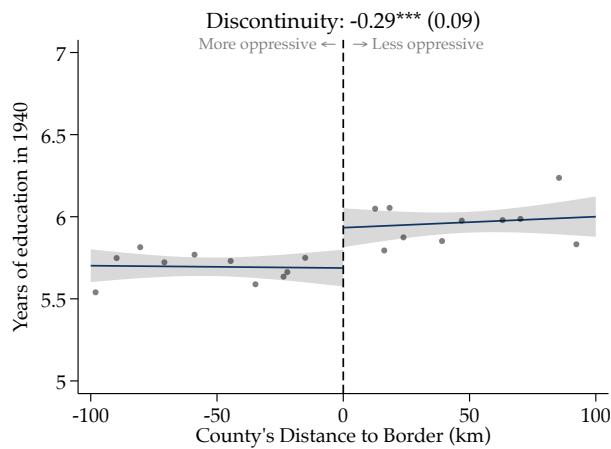
Notes: This figure shows the RD estimate in counties' characteristics in 1860 across state borders with different Jim Crow intensities in 1865. Average income is calculated based on occupational income scores. The sample is restricted to high-contrast borders (0.56 Jim Crow index points or more). The left half of each panel represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of the border population. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands clustered at the county level. See Data Appendix C for details on the sample and data.

FIGURE A.10: Regression Discontinuity Estimates and Education under Jim Crow



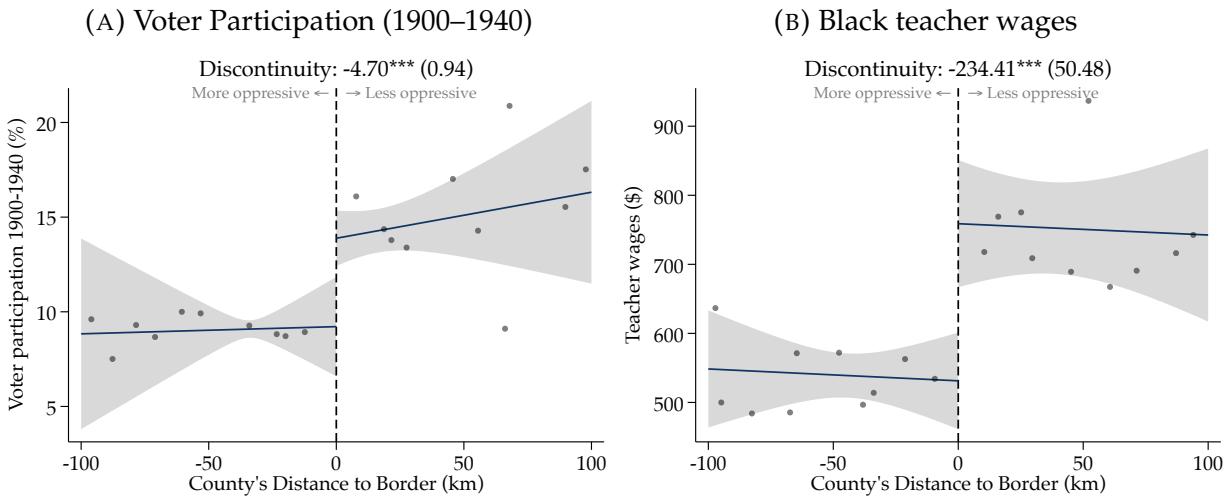
Notes: This figure shows each separate RD estimate in 1940 years of education for Black families whose ancestors were freed on different sides of state borders in 1865. We find qualitatively identical results when using an alternative measure of school quality from Carruthers and Wanamaker (2017) instead of Card and Krueger (1992); results are available upon request. Each label shows the more oppressive before the less oppressive state. Negative estimates reflect lower education in the more oppressive state. Lines show the best linear fit, weighted by the inverse of each estimate's standard error. Shaded areas represent robust 95 percent confidence bands. For point estimates, we use empirical Bayesian shrinkage as described in Appendix B.10. See Data Appendix C for details on the sample and data.

FIGURE A.11: RD Estimates Pooling High- and Low-Contrast Borders



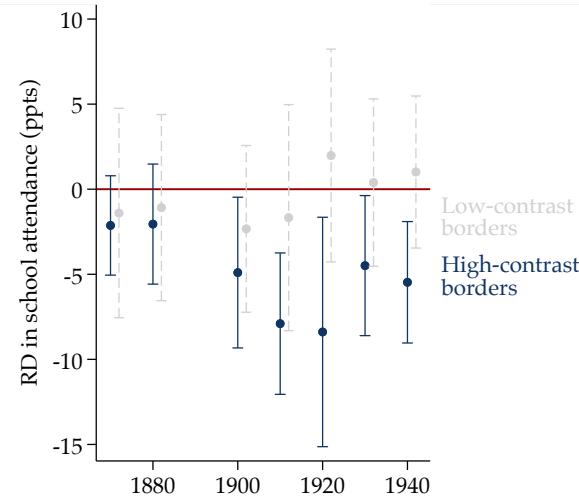
Notes: This figure shows the RD estimate in 1940 years of education for Black families freed across state borders with different Jim Crow intensity in 1865. The left half of the figure represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of the border population. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands clustered at the 1870 county level.

FIGURE A.12: Border Discontinuities in Suffrage and Black Teacher Wages



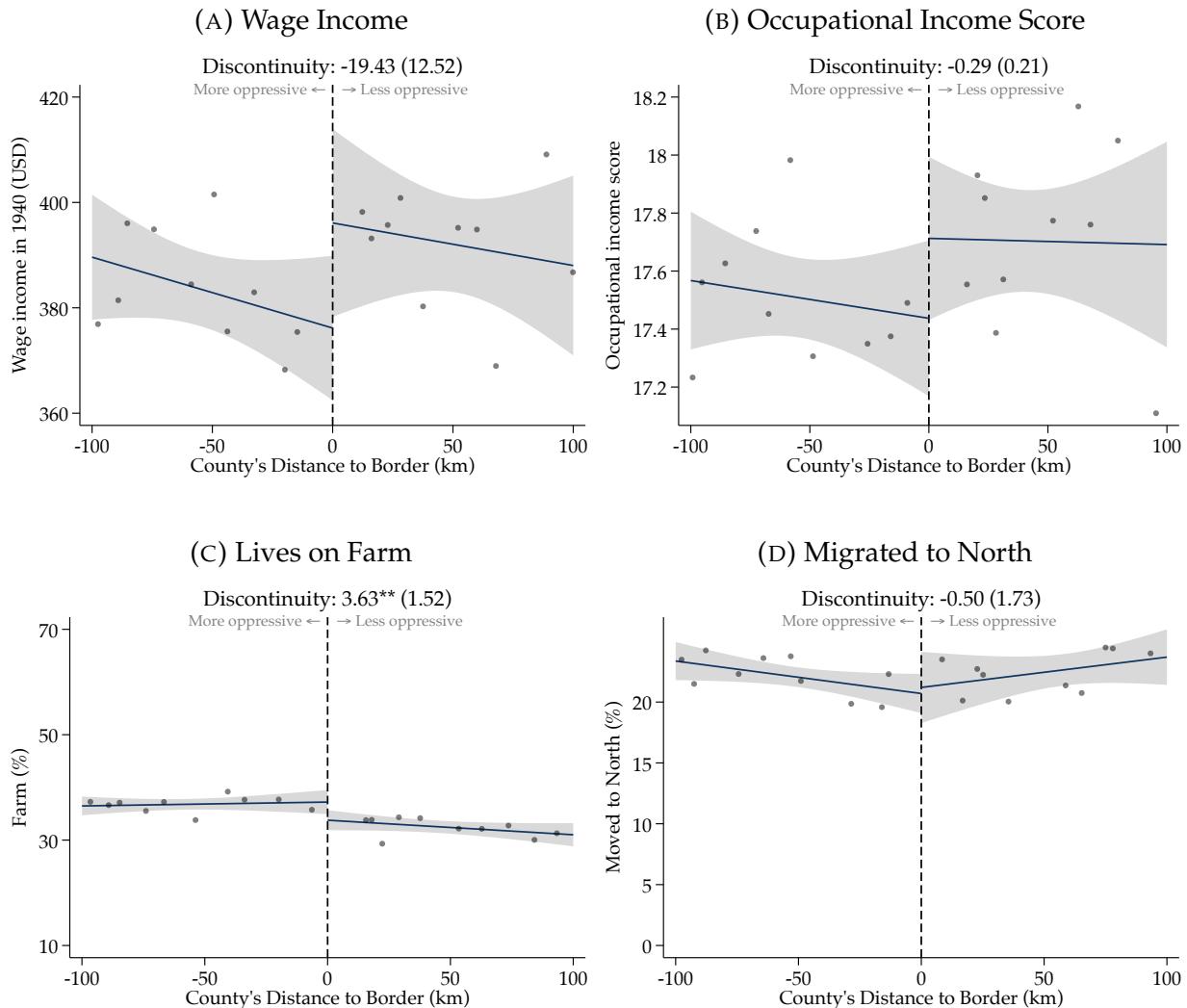
Notes: This figure shows the RD estimates for counties' number of votes cast in decennial Presidential elections from 1900 to 1940 as a share of the total population and counties' Black teacher wages in 1940. The sample is restricted to "high-contrast borders" where Jim Crow intensity differs more than across the median border (0.56 Jim Crow index points or more). The left half of each panel represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of the border population. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands. See Data Appendix C for details on the sample and data.

FIGURE A.13: RDs in Black School Attendance



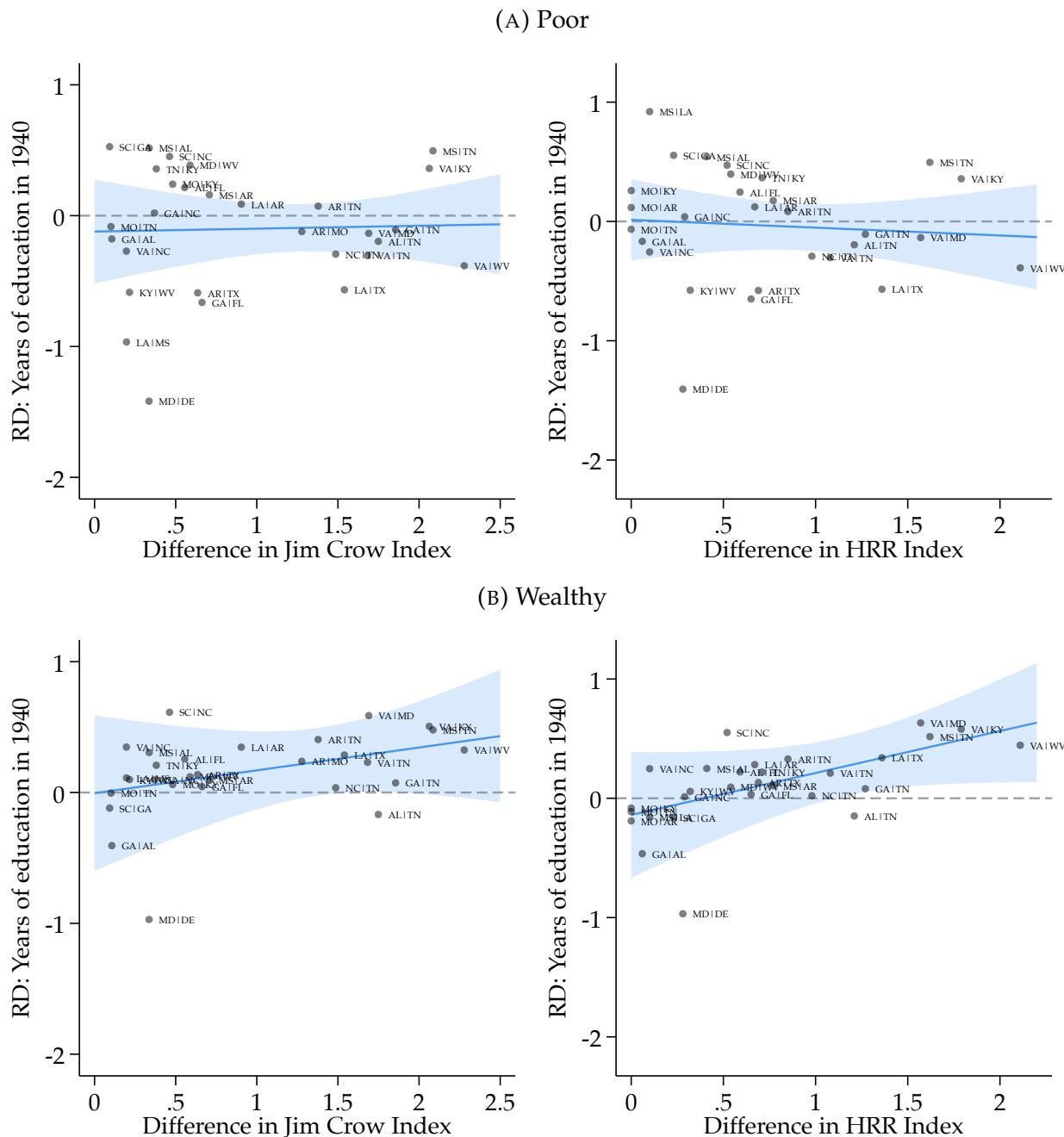
Notes: This figure shows the RD estimates in the school attendance rates of Black children across "high-contrast borders" (0.56 Jim Crow index points or more) and "low-contrast borders." Each estimate is the difference between outcomes in the more oppressive compared to the less oppressive state. Vertical bars represent 95 percent robust confidence bands. See Data Appendix C for details on the sample and data.

FIGURE A.14: Border Discontinuities in Additional 1940 Outcomes



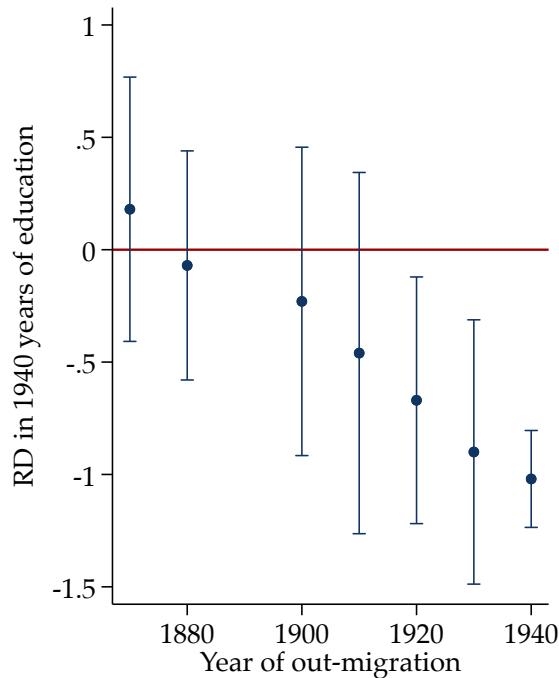
Notes: This figure shows the RD estimate in additional 1940 outcomes for Black families freed across state borders with different Jim Crow intensity in 1865. The sample is restricted to “high-contrast borders” where Jim Crow intensity differs more than across the median border (0.56 Jim Crow index points or more). The left half of each panel represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of the border population. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands clustered at the 1870 county level. See Data Appendix C for details on the sample and data.

FIGURE A.15: RD Estimates for Poor and Wealthy White Americans



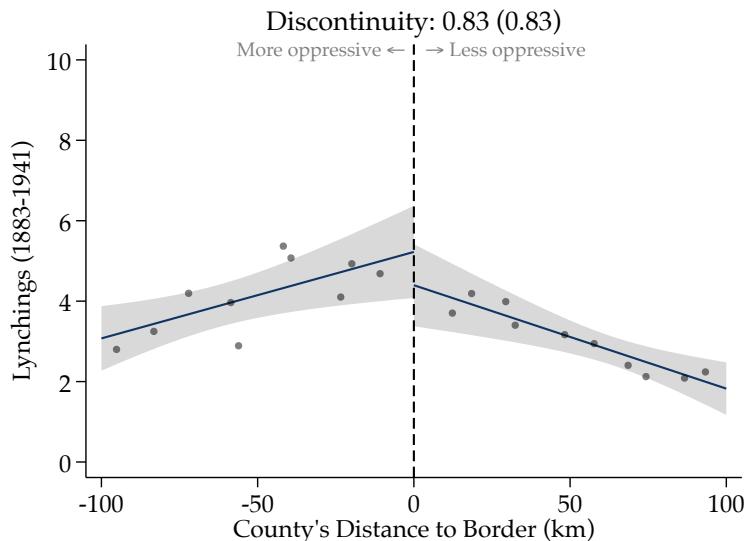
Notes: This figure shows each separate RD estimate in 1940 years of education for white families who had no physical or human capital in 1870, i.e., illiterate and zero wealth (Panel A) or were in the top decile in terms of real property in 1870 (Panel B). Each label shows the more oppressive before the less oppressive state. Negative estimates reflect lower education in the more oppressive state. Lines show the best linear fit between RD estimates and the differences in Jim Crow intensity, weighted by the inverse of each estimate's standard error. Shaded areas represent robust 95 percent confidence bands. For point estimates, we use empirical Bayesian shrinkage as described in Appendix B.10. See Data Appendix C for details on the sample and data.

FIGURE A.16: RD Estimates by Share of Decades Between 1870 and 1940 that a Family Lived in Their Ancestor's Enslavement State



Notes: This figure shows RD estimates in 1940 years of education for Black families whose ancestors were freed on different sides of state borders in 1865 and stayed there for different amounts of time. Each estimate shows the pooled RD estimate for families who stayed in the state where their ancestors were freed from slavery until a given year (x-axis). Negative estimates reflect lower education in the more oppressive state. Bars represent 95 percent confidence intervals. See Data Appendix C for details on the sample and data.

FIGURE A.17: No Border Discontinuities in Lynchings between 1883 and 1941

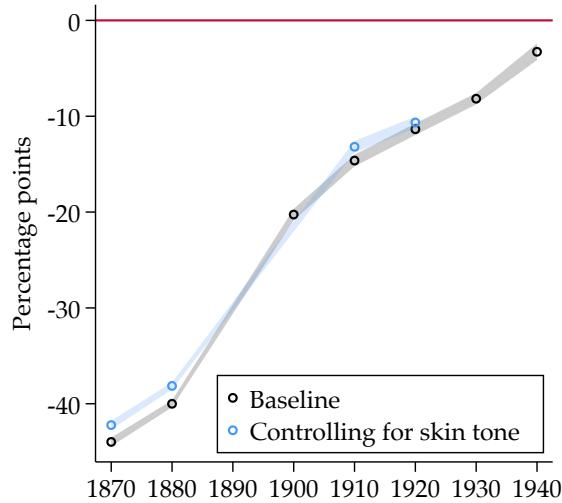


Notes: This figure shows the RD estimate in counties' number of lynchings of Black Americans between 1883 and 1941. The sample is restricted to high-contrast borders (0.56 Jim Crow index points or more). The left half of each panel represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of counties. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands. See Data Appendix C for details on the sample and data.

B. ROBUSTNESS CHECKS

B.1 Free-Enslaved Gap Controlling for Skin-Tone

FIGURE B.18: Free-Enslaved Gap in Literacy Conditional on “Mulatto”-Status



Notes: This figure shows the Free-Enslaved gap in literacy before and after including a dummy for whether a person is classified as “Mulatto” (instead of “Black”) in the census. This classification does not exist in the 1900 census or any census after 1920. The sample includes both the South and North of the US. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. The sample includes only Black prime-age (20–54) men whose ancestors can be located in 1870. See Data Appendix C for details on the sample and data.

B.2 Free-Enslaved Gap Based on the Distribution of Surnames

TABLE B.8: Free-Enslaved Gap Based on the Distribution of Surnames (1940)

PANEL (A): Re-weighted to 1870-level of surnames' relative frequency								
	Education (Years)		Wage Income (USD)		Homeownership (%)		House Value (USD)	
	Mean: 5.70		Mean: 588.60		Mean: 21.53		Mean: 1,616.81	
P(Ancestor Enslaved)	-1.25*** (0.07)	-1.40*** (0.09)	-88.36*** (21.22)	-113.15*** (25.50)	-1.95** (0.87)	-2.31** (1.05)	-1,098.68*** (237.09)	-1,194.53*** (282.83)
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	0.01	0.01	0.01	0.01	0.00	0.00
Observations	2,598,739		2,842,572		2,618,795		556,422	

PANEL (B): Not re-weighted								
	Education (Years)		Wage Income (USD)		Homeownership (%)		House Value (USD)	
	Mean: 5.71		Mean: 598.74		Mean: 21.89		Mean: 1,599.75	
P(Ancestor Enslaved)	-0.47*** (0.02)	-0.54*** (0.02)	-13.73*** (5.26)	-29.89*** (7.17)	-2.43*** (0.21)	-2.43*** (0.29)	-630.53** (277.72)	-970.17* (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 enslaved vs. free Black Americans in 1940. The sample includes both the South and North of the US. 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) surname. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and Enslaved. See Data Appendix C for details on the sample and data. Robust standard errors are shown in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.3 Free-Enslaved Gap Excluding Places with Excess Mortality

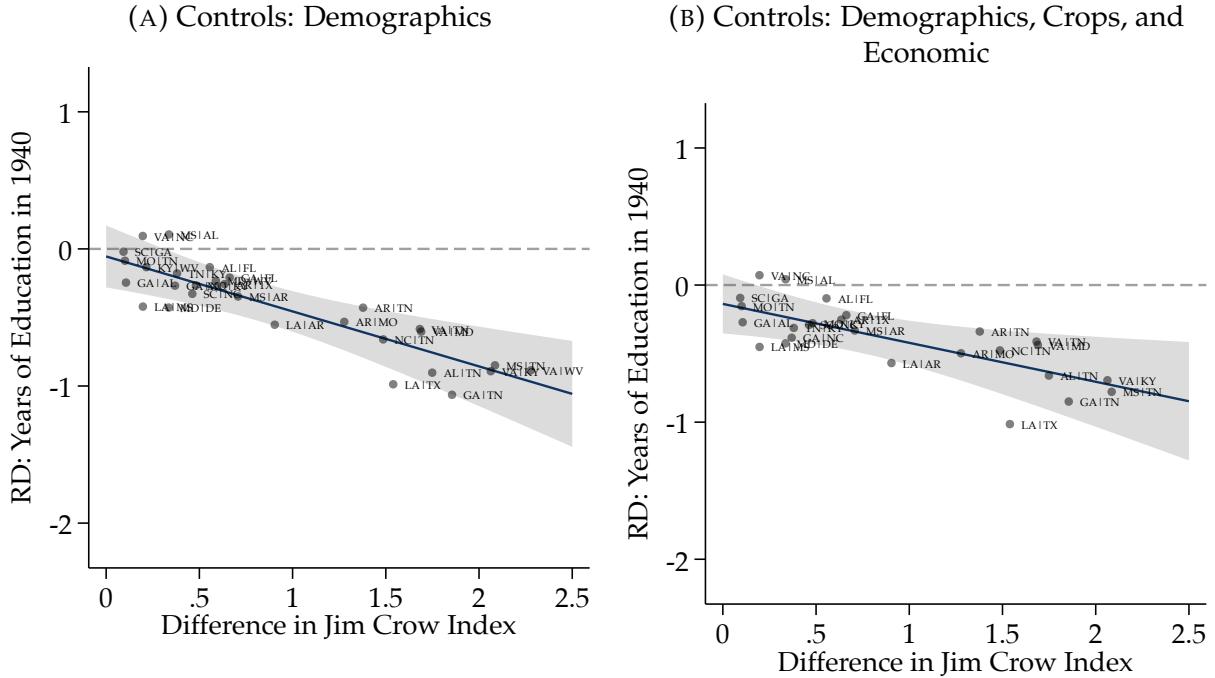
TABLE B.9: Free-Enslaved Gap Excluding Places with Excess Mortality (2000)

	HS Degree (%) Mean: 67.68	College Degree (%) Mean: 11.27	Income (USD) Mean: 27,452.51	House Value (USD) Mean: 93,948.30
Ancestor Enslaved	-3.59*** (0.39)	-2.45*** (0.32)	-4,794.24*** (432.35)	-16,242.52*** (1,862.48)
Level	ZIP × Race × Sex	ZIP × Race × Sex	ZIP × Race	ZIP
Controls (age, age ²)	Y	Y	Y	Y
Adjusted R ²	0.01	0.01	0.01	0.00
Observations	33,932	33,932	33,951	34,274
<i>Ancestor Free</i>	2,196	2,196	2,196	2,219

Notes: This table repeats Table A.3 excluding ZIP codes that have deaths more than twice as high as predicted based on their population density. Each person is assigned the respective value of the ZIP code in which they last lived according to administrative mortality records. The sample includes both the South and North of the US. 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 Enslaved. See Data Appendix C for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.4 Border Discontinuities Controlling for Location Characteristics

FIGURE B.19: RD Estimates Using Different Sets of Control Variables (HRR)

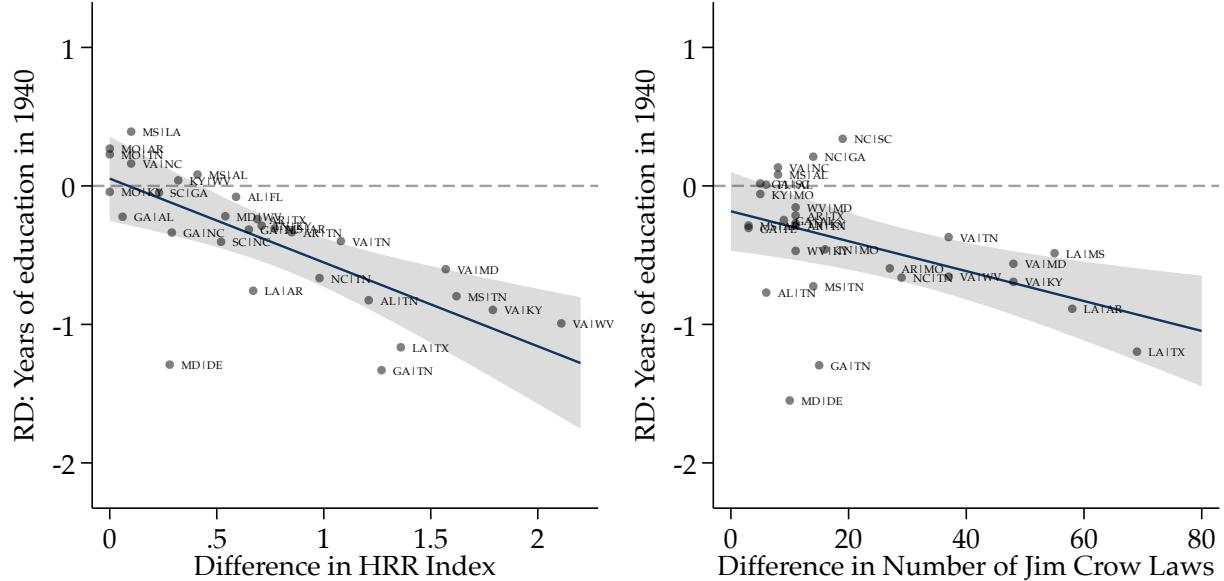


Notes: This figure shows each separate RD estimate in 1940 years of education for Black families freed across state borders with different Jim Crow intensity in 1865 after controlling for different sets of county-level variables in 1860. Panel (A) includes controls for the fraction Black; the fraction free among Black persons; and the age and sex of enslaved persons. Panel (B) includes controls for the farm share; wealth; population density; share Black; migration cost to the North; per-capita tobacco, cotton, and cane sugar output; farm values; and share slaveholders. Each label shows the more oppressive before the less oppressive state. For point estimates, we use empirical Bayesian shrinkage as described in Appendix B.10. See Data Appendix C for details on the sample and data.

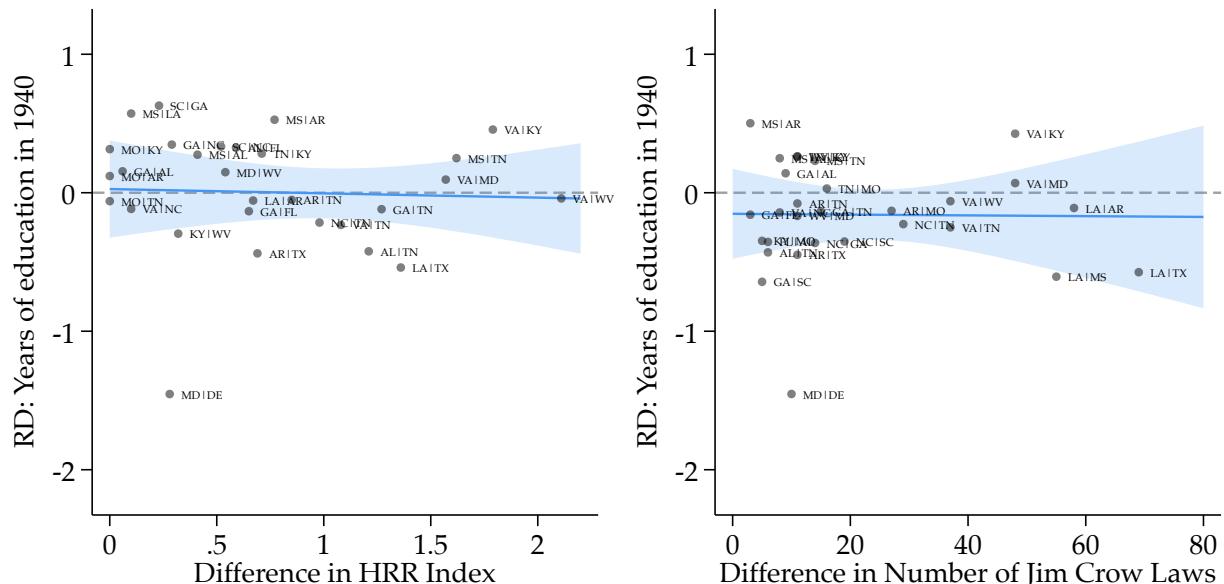
B.5 Border Discontinuities and Alternative Jim Crow Measures

FIGURE B.20: Alternative Jim Crow Intensity Measures and Our RD Estimates

(A) Black Americans



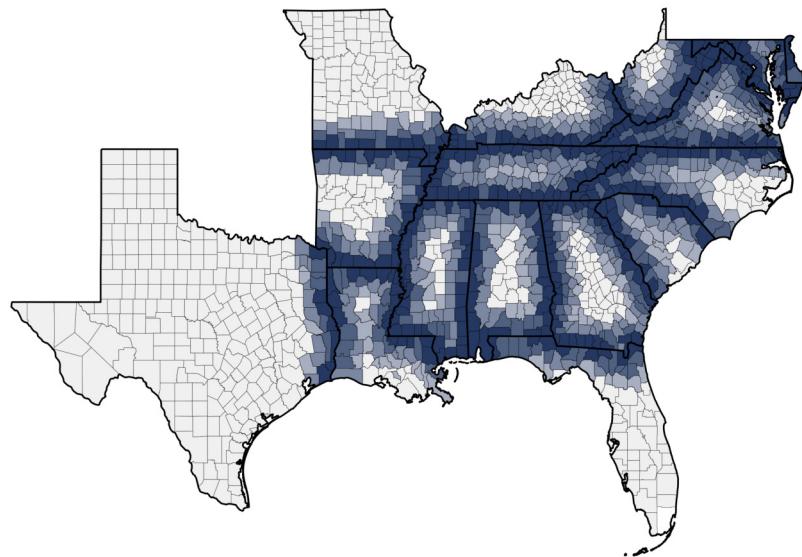
(B) White Americans



Notes: Panel (A) of this figure shows each separate RD estimate in 1940 years of education for Black families whose ancestors were freed on different sides of state borders in 1865. Panel (B) shows the same for white families depending on where their ancestors lived in 1870. Each label shows the more oppressive before the less oppressive state. Negative estimates reflect lower education in more oppressive states. Lines show the best linear fit between RD estimates and the differences in Jim Crow intensity, weighted by the inverse of the estimates' standard error. Shaded areas represent robust 95 percent confidence bands. For point estimates, we use empirical Bayesian shrinkage as described in Appendix B.10. See Data Appendix C for details on the sample and data.

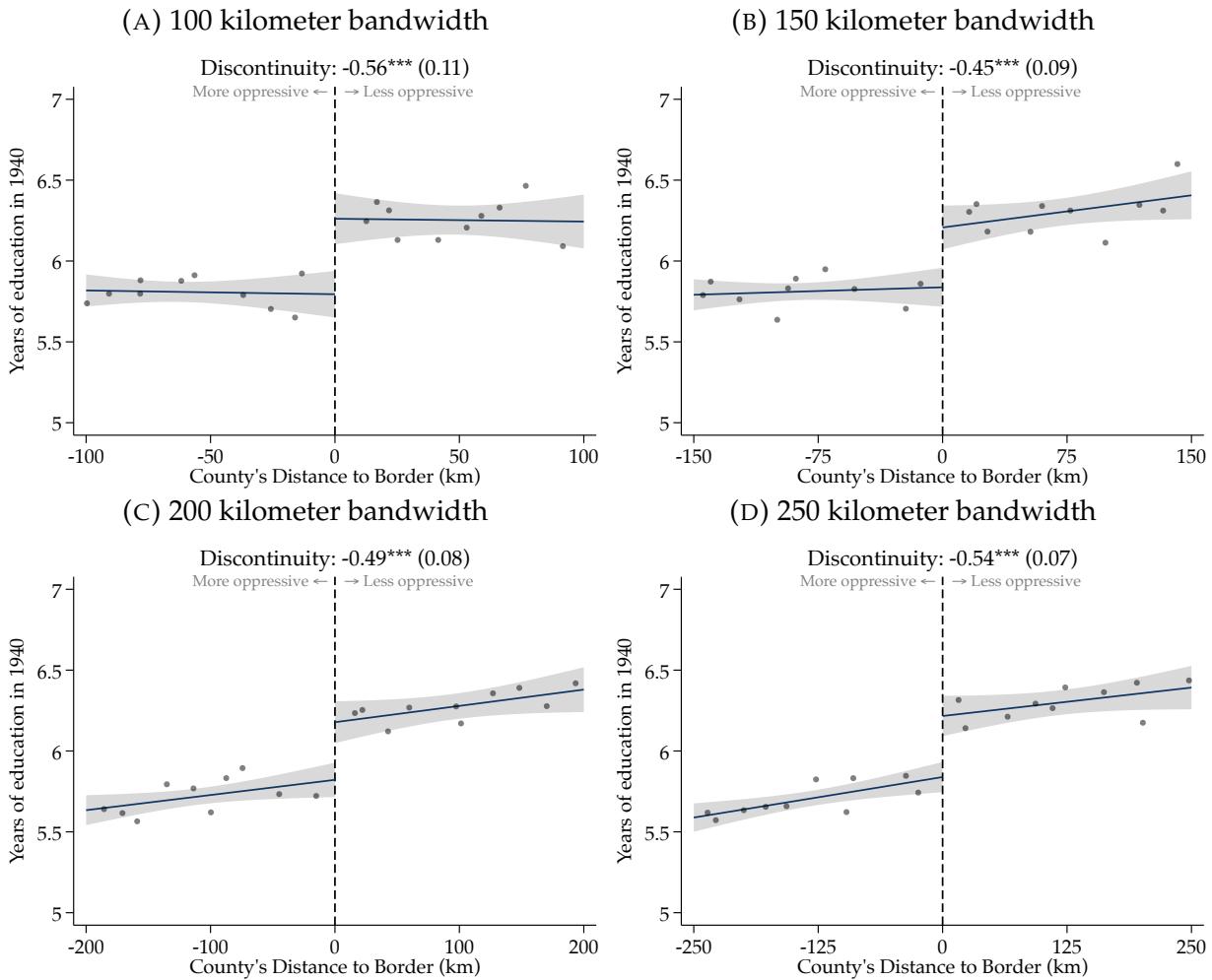
B.6 Border Discontinuities Using Different Bandwidths

FIGURE B.21: Southern Counties' Distance to State Borders



Notes: This map shows each county's distance to the closest state border within the South. Darker shades correspond to closer proximity to a border. Distances are measured from a county's centroid to the border. In our main analysis, we limit our analysis to counties within 100 kilometers (62 miles) of any border but show that our results are robust to other cutoffs.

FIGURE B.22: Different Bandwidths for Pooled RD Estimates



Notes: This figure shows the RD estimate in 1940 years of education for Black families freed across state borders with different Jim Crow intensity in 1865. The analysis is limited to “high-contrast borders” where Jim Crow intensity differs more than across the median border (0.56 Jim Crow index points or more). Panels (A) to (D) show 100, 150, 200, and 250 kilometer bandwidths respectively. The left half of each panel represents more oppressive states; the right half less oppressive states. Each dot is the average across a decile of the border population. Lines show the best linear fit. Shaded areas represent 95 percent confidence bands clustered at the 1870 county level. See Data Appendix C for details on the sample and data.

B.7 Adjusting Estimates for Intermarriage

It is important to distinguishing between two estimands in our analysis: (1) the Free-Enslaved gap based on paternal enslavement ancestry, and (2) the variation in socioeconomic status of a Black individual based on the share of their maternal and paternal ancestors who were Free vs. Enslaved.

The Free-Enslaved gap accurately captures the former estimand, i.e., differences between Black Americans whose male ancestry line goes back to people enslaved until the Civil War vs. Black Americans whose male ancestry line goes back to people free before the Civil War.

The second estimand is more difficult to quantify. Some individuals who we identify as descending from the Free or Enslaved via their paternal ancestry line may descend from the opposite group via other ancestry lines. Because women tended to change their surnames upon marriage, automated linking methods do not allow us to follow a family's female ancestry line directly. However, our estimates of the Free-Enslaved gap can be informative about this second estimand depending on intermarriage levels.

Free-Enslaved intermarriage was likely uncommon, in large part due to the stark geographic and class differences between the two groups. "After the Civil War, the free mulatto class continued to hold itself aloof from the masses of freedmen. In Louisiana, the hostility of some members of this class to the newly emancipated blacks was so great that they opposed giving political rights to the freedmen. [...] Even in their religious affiliations, the descendants of the free mulattoes held aloof from the Negro masses. [...] The descendants of the free mulattoes became, after the Civil War, the core of a small upper class which undertook to maintain the American pattern of family life and conventional sex mores. In some small communities in the South, a single family with this social and cultural background would live in complete isolation rather than associate with the masses of Negroes" ([Frazier, 1957](#)). This historical context suggests that intermarriages across Free-Enslaved status were relatively rare. Below, we provide additional quantitative evidence in support of this view.

For the Free-Enslaved gap, we estimate $y_i = \alpha + \beta \cdot s_i + \varepsilon_i$, while we may also be interested in $y_i = a + b \cdot share_i + e_i$, where $share_i$ is the share of i 's ancestors who were slave until the Civil War. For our estimate of the Free-enslaved gap, we have

$$\hat{\beta} \xrightarrow{p} \mathbb{E}[y|s=1] - \mathbb{E}[y|s=0] = b \cdot (\mathbb{E}[share_i|s=1] - \mathbb{E}[share_i|s=0]). \quad (13)$$

We use this expression to derive the attenuation bias that makes the Free-Enslaved gap a lower bound for the true group differences between families with high vs. low shares of ancestors enslaved.

REFERENCES

FRAZIER, E. F. (1957): *Black bourgeoisie*, Free Press.

B.7.1 First generation after slavery

For the first generation of descendants, we know that

$$\begin{aligned}\mathbb{E}[share_{i,1}|s=1] &= 1 \cdot \mathbb{P}(share_{i,1} = 1|s_i = 1) + 0.5 \cdot \mathbb{P}(share_{i,1} = 0.5|s_i = 1) + 0 \\ &= 1 \cdot \mathbb{P}(\text{mother slave}|\text{father slave}) + 0.5 \cdot \mathbb{P}(\text{mother free}|\text{father slave}) \\ \mathbb{E}[share_{i,1}|s=0] &= 1 \cdot \mathbb{P}(share_{i,1} = 1|s_i = 0) + 0.5 \cdot \mathbb{P}(share_{i,1} = 0.5|s_i = 0) + 0 \\ &= 0.5 \cdot \mathbb{P}(\text{mother slave}|\text{father free})\end{aligned}$$

Therefore, we have

$$\hat{\beta} \xrightarrow{p} b_1 \cdot [0.5 + 0.5 \cdot \mathbb{P}(\text{mother slave}|\text{father slave}) - 0.5 \cdot \mathbb{P}(\text{mother slave}|\text{father free})].$$

If there was no intermarriage, we would have $\hat{\beta} \xrightarrow{p} b_1$.³¹ If marriage between formerly enslaved families and free Black families were random—in the sense that free and enslaved fathers have an equal probability of marrying an enslaved mother—we would have $\hat{\beta} \xrightarrow{p} 0.5 \cdot b_1$.³² Given that it is implausible that free Black men were more likely than formerly enslaved Black men to marry formerly enslaved women, it seems reasonable that $b_1 \in [\hat{\beta}, 2 \cdot \hat{\beta}]$.

We empirically assess this bias by analyzing the likelihood that a Black person descends from one parent born in a slave state and another parent born in a free state for 20-40 year old Americans in the 1910 census (whose parents were likely born towards the end of slavery). We are not able to quantify intermarriage between the formerly Enslaved and Free within state of origin because we do not have information on women's enslavement status beyond her birthplace.

We estimate that in 1910,

$$\begin{aligned}\hat{\mathbb{P}}(\text{mother slave}|\text{father slave}) &= 0.99 \\ \hat{\mathbb{P}}(\text{mother slave}|\text{father free}) &= 0.20,\end{aligned}$$

suggesting that the gap between individuals whose grandparents are either all formerly Enslaved or all Free could be 1.1 times as large as the Free-Enslaved gap.

Our state-based proxy for intermarriage may deviate from actual rates for two reasons. First, our proxy could underestimate actual rates by ignoring Free-Enslaved marriages among those born in the same region (South). However, the historical accounts about deep class divides make frequent intermarriages unlikely even for Black Americans born within the same location. Second, our state-based proxy could overstate actual rates of

³¹Without intermarriage: $\mathbb{P}(\text{mother slave}|\text{father slave}) = 1$ and $\mathbb{P}(\text{mother slave}|\text{father free}) = 0$.

³²With random intermarriage: $\mathbb{P}(\text{mother slave}|\text{father free}) = \mathbb{P}(\text{mother slave}|\text{father slave}) = \mathbb{P}(\text{mother slave})$.

Free-Enslaved marriages because, as we show in this paper, the Free were more likely to migrate from the South to the North than the Enslaved, inflating the number of Free-Free marriages we incorrectly classify as Free-Enslaved marriages.

B.7.2 Second generation after slavery

If there was no intermarriage, we would have $\hat{\beta} \xrightarrow{p} b_2$. If marriage between formerly enslaved families and free Black families were random we would have $\hat{\beta} \xrightarrow{p} 0.25 \cdot b_2$. Thus, $b_2 \in [\hat{\beta}, 4 \cdot \hat{\beta}]$. The details of the derivation are available upon request.

We empirically assess this bias by analyzing the likelihood of having parents born in slave or free states for married couples between 20 and 40 years old in the 1910 census (whose parents were likely born towards the end of slavery). Our estimates suggest that the gap between individuals whose grandparents are either all formerly Enslaved or all Free could be 1.5 times as large as the Free-Enslaved gap.

B.7.3 n^{th} generation after slavery

Generally, if there was no intermarriage, we would have $\hat{\beta} \xrightarrow{p} b_n$. If marriage between formerly enslaved families and free Black families were random we would have $\hat{\beta} \xrightarrow{p} 2^{-n} \cdot b_n$. Thus, $b_n \in [\hat{\beta}, 2^n \cdot \hat{\beta}]$.

Our geographic ancestry analysis from 1880 to 1940 indicates little intermarriage between slave and non-slave states even in the latest decades of our sample period. Specifically, the probability of a Black person's mother being born in a slave state, given their father was also born in a slave state, is between 98 and 100 percent throughout this period. Conversely, for fathers born in free states, the probability that the mother was also from a free state ranges between 64 and 86 percent (while free Black Americans only account for 11 percent of the Black population).

B.8 Adjusting Estimates for Misclassification Bias

Potential misclassification of ancestors' enslavement status could bias our estimates of the Free-Enslaved gap towards zero. It is valuable to distinguish two kinds of misclassification: false negatives, which refer to individuals incorrectly classified as formerly Enslaved despite having free paternal ancestry (due to imperfect linking rates); and false positives, which refer to individuals incorrectly classified as Free when their paternal ancestry was enslaved until the Civil War (due to incorrect links to the 1850 or 1860 census).

To mitigate misclassification bias, we use an instrumental variable (IV) approach designed to correct for both false negatives and false positives. We use our surname-based measure as an instrument for the linking-based measure. The resulting IV estimates of

fer an unbiased assessment of the Free-Enslaved gap, contingent upon the measurement errors in the linking-based measure being uncorrelated with the surname-based measure ([Ashenfelter and Krueger, 1994](#); [Angrist and Pischke, 2008](#)). This assumption is plausible given that our surname-based measure is independent of census-linking methods.

The IV results suggest that measurement error reduces our initial estimates of the Free-Enslaved gap by an average of 9 percent across various outcomes (see Table B.10). For instance, the education gap, as estimated via the IV approach, stands at 1.67 years—a 5 percent increase compared to the OLS estimate of 1.59 years.

TABLE B.10: Free-Enslaved Gap (1940): IV Approach to Reduce Measurement Error in Enslavement Status

	Education (Years)	Wage Income (USD)	Home Ownership (%)	House Value (USD)
	Mean: 6.08	Mean: 390.18	Mean: 29.71	Mean: 1,422.37
IV: Ancestor Enslaved	-1.67*** (0.15)	-170.12*** (17.69)	-9.69*** (1.89)	-554.68*** (149.68)
OLS: Ancestor Enslaved	-1.59***	-145.92***	-7.24***	-694.69***
Controls (age, age ²)	Y	Y	Y	Y
F-Statistic (weak id.)	2,077.22	1,998.63	2,049.38	994.86
Adjusted R ²	0.05	0.05	0.01	0.01
Observations	158,032	149,252	158,787	45,311
<i>Ancestor Free</i>	9,078	8,551	9,070	3,227

Notes: This table shows instrumental variable (IV) estimates of the gap in years of education, wage income, homeownership, and house value (conditional on ownership) among prime-age (20–54) male descendants of enslaved vs. free Black Americans in 1940. We use our surname-based measure of a Free-Enslaved status as an instrument for our linking-based measure. The sample includes both the South and North of the US. Only observations that can be linked to the 1850, 1860, 1870, or 1880 census are included. Sample means are computed for the combined sample of the Free and Enslaved. See Data Appendix C for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We also separately address potential bias from false negatives, which is more likely to be significant due to the conservative nature of our linking approach that makes false positives unlikely. The linking criteria require both uniqueness within and matches across two census waves, based on several attributes including name, year and state of birth, and race. Our methodology may incorrectly categorize many Black families as descendants of the Enslaved, particularly if they originated in slave states with a significant pre-Civil War free Black population. For instance, in Maryland, approximately 50 percent of Black Americans were free before the Civil War according to the 1860 census. In our sample, 70 percent of Black Americans with ancestors from Maryland are classified as descendants of the Enslaved in 1940—20 points more than expected.

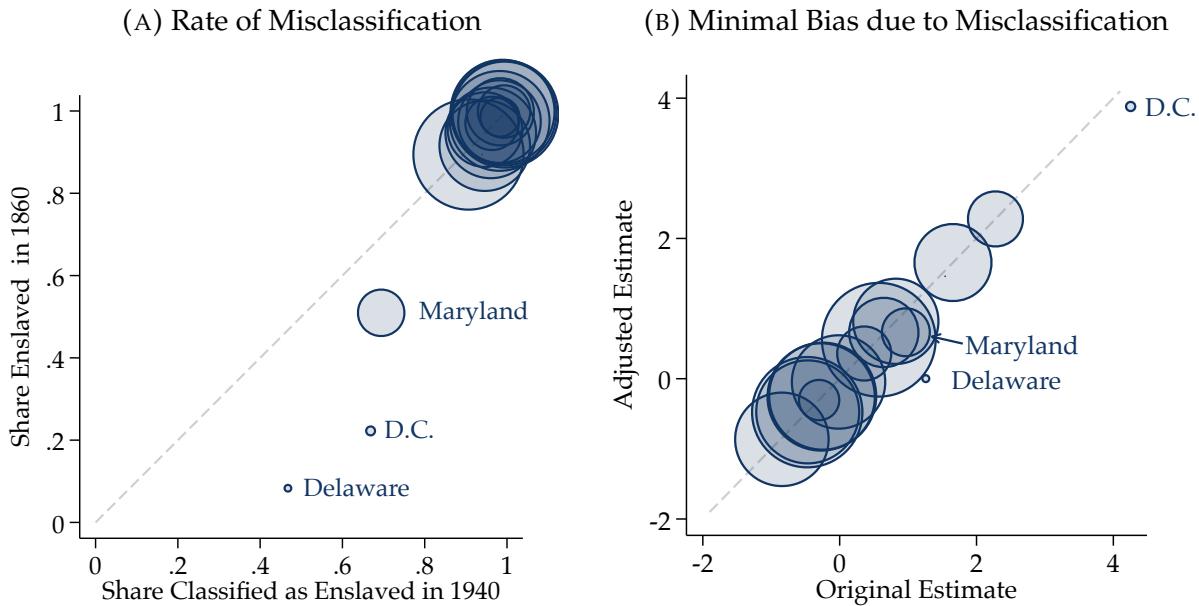
We adjust our estimates for bias that may arise from this type of misclassification. We use that our original estimates are a weighted average of the (unknown) unbiased

estimate and the non-causal estimate for free Black Americans:

$$\hat{\beta}_{\text{original}} = \frac{\text{Enslaved}_{s,\text{links}}}{\text{Enslaved}_{s,1860}} \cdot \hat{\beta}_{\text{unbiased}} + \left(1 - \frac{\text{Enslaved}_{s,\text{links}}}{\text{Enslaved}_{s,1860}}\right) \cdot \hat{\beta}_{\text{free}}, \quad (14)$$

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 $\hat{\beta}_{\text{free}}$ is the non-causal estimate for outcomes of those with ancestors from state s .

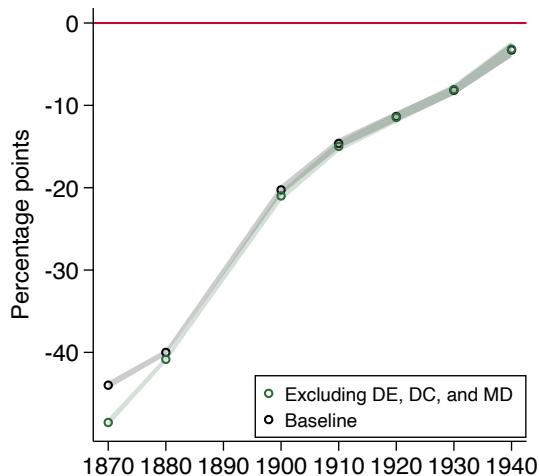
FIGURE B.23: Misclassification and Bias



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. The sample includes the South of the US. See Data Appendix C for details on the sample and data.

We find that adjusting for the gap between the actual proportion of free Black individuals before the Civil War and our smaller classified share has a small impact on our Free-Enslaved gap estimates. Figure B.23 shows that the share of Black Americans who descend from the Enslaved only deviates from our classification for three small slave states. Accordingly, adjusting our original estimates of the causal effect of each state barely affects our estimates. Even when excluding states with a high pre-Civil War free Black population, our gap estimate remains largely unchanged (see Figure B.24).

FIGURE B.24: Free-Enslaved Gap in Literacy (1870–1940)



Notes: This figure shows the gaps in literacy among prime-age (20-54) male descendants of enslaved vs. free Black Americans in each census decade before and after excluding Delaware, DC, and Maryland. The sample includes both the South and North of the US. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. 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. See Data Appendix C for details on the sample and data.

REFERENCES

- ANGRIST, J. D. AND J.-S. PISCHKE (2008): *Mostly harmless econometrics*, Princeton university press.
- ASHENFELTER, O. AND A. KRUEGER (1994): “Estimates of the Economic Return to Schooling from a New Sample of Twins,” *The American Economic Review*, 84, 1157–1173.

B.9 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. Under a mild assumption, we can recover the treatment effect of each destination location.

Assumption 2 (No direct long-run effect of enslavement location). *The pre-1865 effect of enslavement location ℓ ceases to directly affect a family's descendants by 1940. That is,*

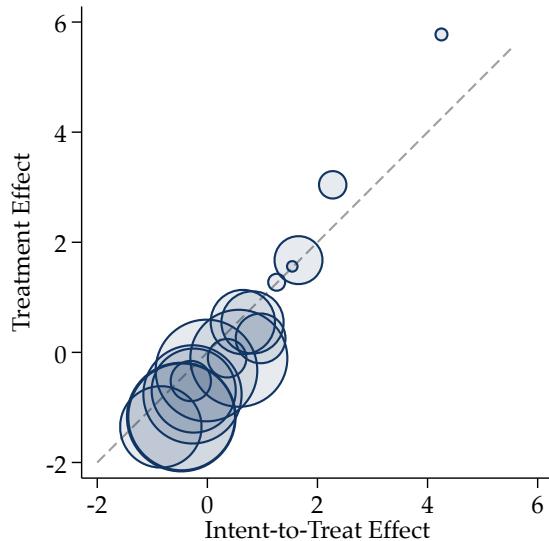
$$\rho\gamma_c^0 = 0$$

where ρ is the intergenerational elasticity from 1865 to 1940 and γ_ℓ^0 is the effect that location ℓ had on Black families who lived there.

This assumption is plausible for two reasons. First, the vast majority of enslaved people were freed from slavery with little to no measurable physical or human capital with little variation across locations. Second, plausible values for ρ are likely small given the high intergenerational mobility of Black Americans following the end of slavery and the amount of time that elapsed until 1940.³³

Under this assumption, we can recover a state's treatment effect from the originally estimated intent-to-treat (ITT) using standard instrumental variable methods in settings with multiple treatments under imperfect compliance—each treatment being a potential state of birth and non-compliance arising through migration. As described in Section 6.1, the ITT effect of location ℓ , η_ℓ , is the average of all potential future locations' treatment effects, $\gamma_{\ell'}^1$, weighted by the probability of migrating from ℓ to ℓ' . We invert the migration probability matrix to recover the effect of living in each state until 1940.

FIGURE B.25: ITT Effect and Treatment Effect 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 treatment effect that living in that state had. The estimates are in years of education in 1940. See Data Appendix C for details on the sample and data.

We find that the original ITT effect of living in a state after 1865, estimated as the causal effect of being born into slavery in that state, is almost identical to the treatment effect of living in the state after 1865 (see Figure B.25). In essence, this finding results from high “compliance rates” due to limited geographic mobility in the Deep South before 1940.

³³With an intergenerational elasticity of 0.25 for Black Americans (Althoff et al., 2023) and around three generations elapsing between 1865 and 1940: $\rho = 0.25^3 \approx 0.02$.

REFERENCES

ALTHOFF, L., H. BROOKES GRAY, AND H. REICHARDT (2023): “The Missing Link(s): Women and Intergenerational Mobility,” Work in progress.

B.10 Empirical Bayes Shrinkage

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 form 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.

We use an 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.

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

$$\hat{\theta}_c = X'_c \hat{\delta} + \frac{\hat{s}}{\hat{s} + \hat{\sigma}_c^2} \cdot (\eta_c - X'_c \hat{\delta}),$$

where $\hat{\eta}_c$ is the preliminary estimate of county c 's effect,

$$\hat{\delta} = \left(\sum_{c=1}^N \hat{\sigma}_c^{-2} X_c X'_c \right)^{-1} \sum_{c=1}^N \hat{\sigma}_c^{-2} X_c \hat{\eta}_c$$

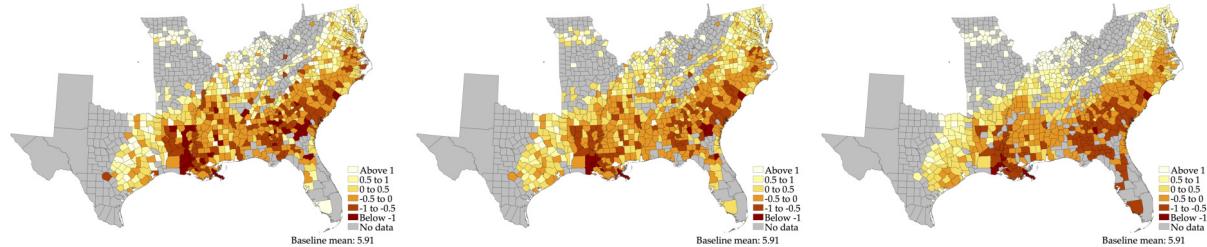
is the ordinary least squares estimate of $\hat{\eta}_c$ on the county covariates X_c , $\hat{\sigma}_c^2$ is the standard error of $\hat{\eta}_c$, and $\hat{s} = \max \left\{ \frac{-N + \sum_{c=1}^N \hat{\sigma}_c^{-2} \hat{\varepsilon}_c^2}{\sum_{c=1}^N \hat{\sigma}_c^{-2}}, \frac{2}{\sum_{c=1}^N \hat{\sigma}_c^{-2}} \right\}$ with $\hat{\varepsilon}_c = \hat{\eta}_c - X'_c \hat{\delta}$. 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 more precisely estimated preliminary effects.

Figure [B.26](#) shows the place effects before and after shrinkage. While the negative effects are concentrated in the Lower South before and after, the shrunk estimates are more spatially correlated. Figure [B.27](#) shows the correlation of causal place effects on Black

economic progress with the same places' (non-causal) effects on the outcomes of white and free Black Americans. Before and after shrinkage, there is no correlation between the effects for descendants of the Enslaved and white Americans, but a strong positive correlation between those for descendants of the Enslaved and the Free.

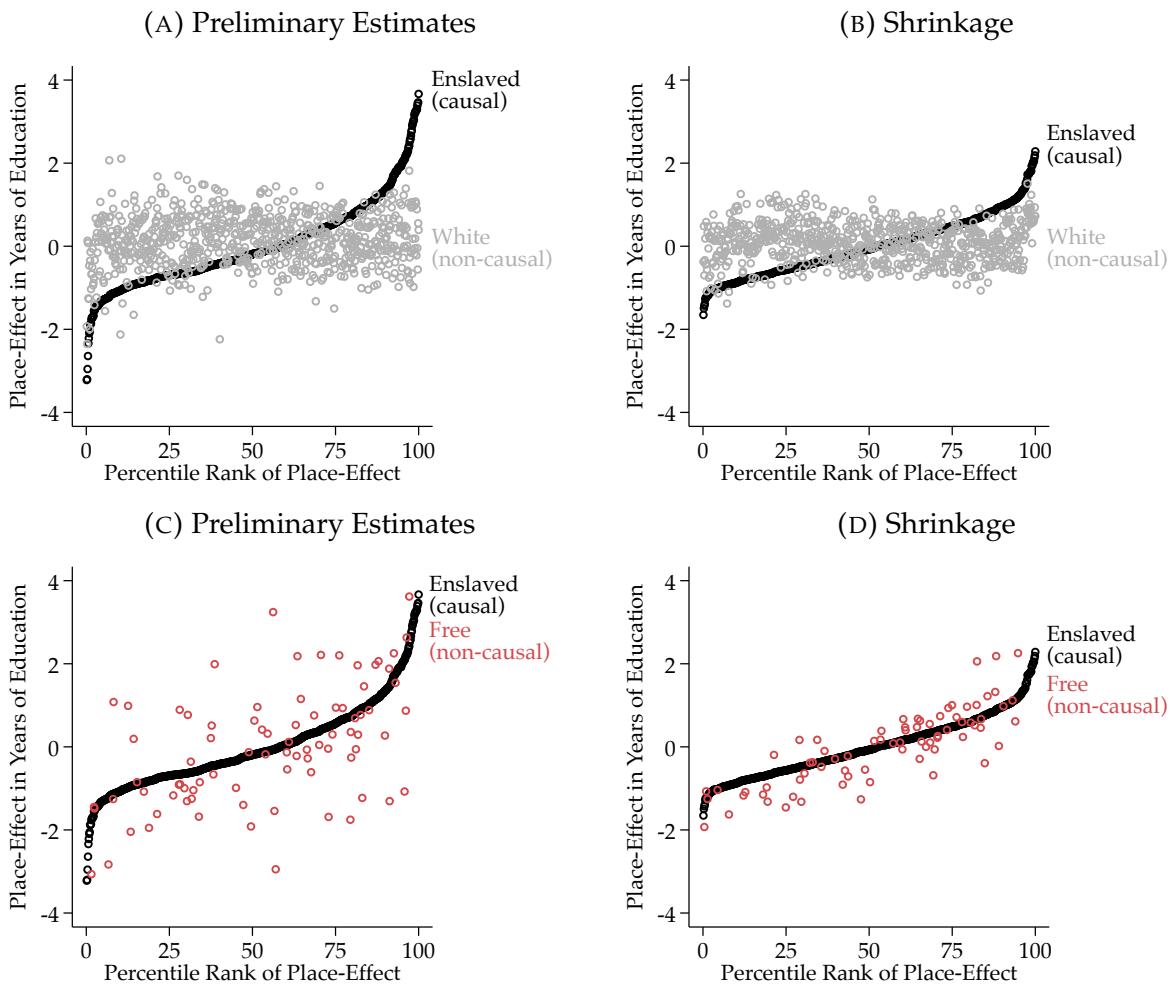
FIGURE B.26: Causal Place Effects on 1940 Years of Education

(A) Preliminary Estimates (B) Shrinkage (No Covariates) (C) Shrinkage (Covariates)



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 Data Appendix C for details on the sample and data.

FIGURE B.27: 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 (non-causal). Panels (A) and (C) show the estimates before shrinkage, Panels (B) and (D) show the shrinkage estimates. The shrinkage does not preserve a county's original rank. County-fixed effects based on ten observations or fewer are discarded. See Data Appendix C for details on the sample and data.

REFERENCES

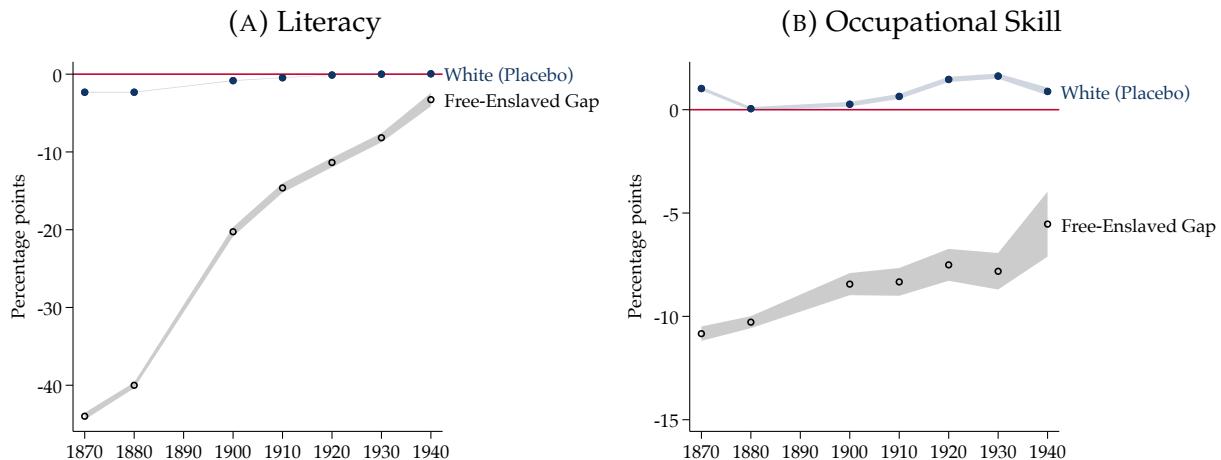
- ANGRIST, J. D., P. D. HULL, P. A. PATHAK, AND C. R. WALTERS (2017): "Leveraging Lotteries for School Value-Added: Testing and Estimation," *Quarterly Journal of Economics*, 132, 871–919.
- ARMSTRONG, T. B., M. KOLES, AND PLAGBORG-MØLLER (2021): "Robust Empirical Bayes Confidence Intervals," Tech. 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.

B.11 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.³⁴ When evaluating the placebo estimates, we should bare in mind that these changes contaminate a pure placebo.

FIGURE B.28: Free-Enslaved Gap (1870-1940) vs. Placebo for White Americans



Notes: This figure shows the true and placebo gaps in literacy rates and occupation skill levels among prime-age (20-54) male descendants of enslaved vs. free 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. The sample includes both the South and North of the US. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. We assign “skilled” to occupations classified as “medium skilled workers” or above by the HISCLASS scheme (Leeuwen and Maas, 2011); and “unskilled” to others. 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. See Data Appendix C for details on the sample and data.

³⁴Indeed, we find that the composition of white Americans indeed experienced some changes after 1860. White Americans grow more likely to be first-generation immigrants and among those immigrants fewer come from the United Kingdom and more from Northern, Central, and Eastern Europe. Results available upon request.

TABLE B.11: Placebo Free-Enslaved Gap (1940) for White Americans

	Education (Years)	Wage Income (USD)	Home Ownership (%)	House Value (USD)
	Mean: 9.76	Mean: 892.68	Mean: 49.74	Mean: 3,284.56
Placebo	-0.17*** (0.00)	-1.68 (1.04)	0.09 (0.05)	12.17 (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
<i>Ancestor Free</i>	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. The sample includes both the South and North of the US. 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 Enslaved. See Data Appendix C for details on the sample and data. Standard errors are clustered at the family level and are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Second, we estimate the Free-Enslaved gap using 1875 as the (placebo) end of slavery. Figure 2 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 surnames from 1870 to 1880.

TABLE B.12: Free-Enslaved vs. Placebo Gap using Surname-Based Approach (1940)

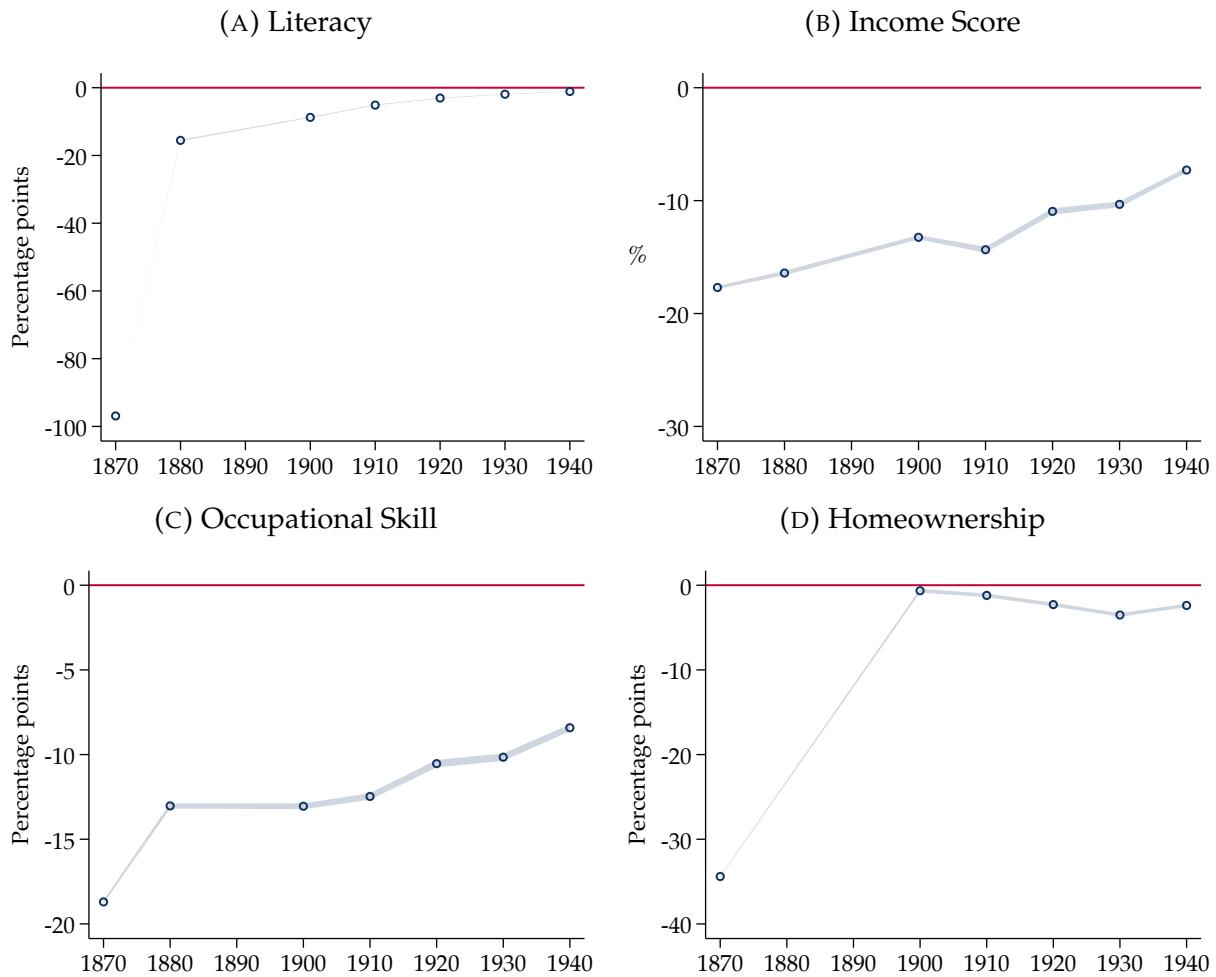
	Education (Years)			
P(Ancestor Enslaved)		-0.47*** (0.02)	-0.43*** (0.02)	-0.54*** (0.02)
Placebo: 1875	-0.05*** (0.01)	-0.03*** (0.01)		0.21*** (0.01)
Placebo: White		-0.04*** (0.00)	-0.04*** (0.00)	0.07*** (0.00)
Placebo: 1875×White				-0.24*** (0.00)
Name-measure	Exact	Exact	Exact	Exact
Controls (race, race×age, race×age ²)	Y	Y	Y	Y
Adjusted R ²	0.03	0.03	0.03	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 enslaved vs. free Black Americans in 1940. The sample includes both the South and North of the US. 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) surname. House values are measured conditional on ownership. Sample means are computed for the combined sample of the Free and Enslaved. See Data Appendix C for details on the sample and data. Robust standard errors are shown in parentheses.
 *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B.12 Convergence Benchmark

Beyond the above mentioned 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 do not have a causal interpretation. 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.

FIGURE B.29: 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. In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. We assign “skilled” to occupations classified as “medium skilled workers” or above by the HISCLASS scheme ([Leeuwen and Maas, 2011](#)); and “unskilled” to others. 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. See Data Appendix C for details on the sample and data.

C. DATA APPENDIX

C.1 Jim Crow Database

We build a rich dataset on states' Jim Crow regimes by combining newly collected information on Jim Crow laws and existing data on states' institutions and outcomes directly affected by those institutions, including voter participation and educational resources.

C.1.1 Jim Crow Index

To measure the intensity of each state's Jim Crow regime, we introduce a composite metric—the “Jim Crow index.” This index is constructed using principal component analysis and encompasses multiple factors, each serving as a proxy for specific aspects of anti-Black legislation. Our index builds on the “Historical Racial Regime (HRR) index” from [Baker \(2022\)](#) but focuses on institutional factors and the Jim Crow era specifically.

Our new Jim Crow index is based on five factors. The first factor is the anti-Black share of race-related laws a state passed until 1950. For this measure, we collected new information on laws that mention race or color and classify those laws as to whether they are anti-Black discriminatory or not (see next section). The second factor is a state's number of disenfranchisement devices (i.e., literary tests, poll tax, grandfather clause, and White primary; [Walton et al., 2012](#); [Baker, 2022](#)). The third factor is a state's share of congressional delegates that signed the Southern Manifesto ([sou, 1956](#); [Baker, 2022](#)). The fourth factor is the racial gap in states' school year lengths—i.e., the legislative term length of Black schools relative to the legislative term length of white schools ([Card and Krueger, 1992](#)). The fifth and final factor is the year in which a state introduced legislation for minimum teacher pay—legislation central to narrowing the large wage penalty historically suffered by Black teachers ([Card et al., 2022](#); [Cascio and Lewis, 2022](#)).

Table C.13 presents each state's Jim Crow Index alongside the corresponding input variables. The “Deep South” states—Mississippi, Louisiana, Georgia, South Carolina, and Alabama—emerge as the most oppressive according to our index. Notably, Louisiana ranks in the top quartile of most oppressive states across all measures. In contrast, the border states—Delaware, West Virginia, Kentucky, Maryland, and Missouri—are categorized as the least oppressive.

TABLE C.13: The Jim Crow Index

State	Jim Crow Index	Share of laws discriminatory	Disenfranchisement devices	Southern Manifesto	Black-white ratio in term length	Minimum teacher pay introduced
Louisiana	1.33	96%	4	100%	0.77	1948
Mississippi	1.14	98%	3	100%	0.78	1924
South Carolina	1.00	92%	3	100%	0.76	1945
Georgia	0.91	96%	4	100%	0.91	1937
Alabama	0.80	93%	4	100%	0.89	1927
Virginia	0.73	93%	4	100%	0.95	1946
North Carolina	0.54	96%	4	71%	0.96	1919
Arkansas	0.43	88%	2	100%	0.88	1957
Florida	0.24	92%	2	80%	0.96	1955
Texas	-0.21	89%	2	21%	0.93	1949
Missouri	-0.85	88%	0	0%	1.05	1985
Tennessee	-0.95	80%	1	36%	0.99	1925
Maryland	-0.96	89%	0	0%	0.96	1904
Delaware	-1.29	82%	0	0%	1.00	1919
Kentucky	-1.33	85%	0	0%	1.05	1912
West Virginia	-1.54	81%	0	0%	1.00	1882

Notes: This table shows each states' Jim Crow Index, ordered from most to least oppressive. The Jim Crow Index is a principle component extracted from five factors, as shown in the remaining columns. The **top-quartile (most oppressive)** is highlighted in red; the **bottom-quartile (least oppressive)** in blue.

We consider a variety of alternative measures for states' Jim Crow intensity. Figure C.30 shows the correlations between different proxies of Jim Crow intensity (discussed in the following two sections). While these measures are very different in nature and capture both de jure and de facto aspects of Jim Crow, they are correlated and using them, we consistently arrive at the same conclusions. Key outcomes directly affected by Jim Crow institutions are also highly correlated with our Jim Crow index: overall voter participation between 1900 and 1940 ($\rho = -0.89$, not available by race) and our causal effects on long-run economic progress of Black families ($\rho = -0.93$).

FIGURE C.30: Correlations Between Proxies of Jim Crow Intensity

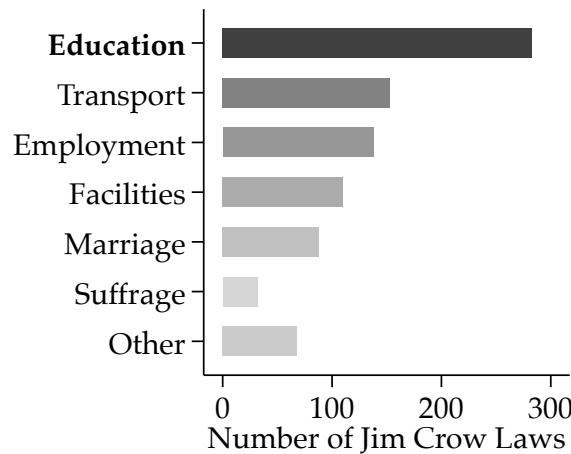
Jim Crow index	1.00	0.99	0.77	-0.89	-0.91	-0.93
HRR index	0.99	1.00	0.74	-0.92	-0.94	-0.92
Jim Crow laws	0.77	0.74	1.00	-0.64	-0.57	-0.69
Voter participation	-0.89	-0.92	-0.64	1.00	0.92	0.82
Black school quality	-0.91	-0.94	-0.57	0.92	1.00	0.85
Causal state effect	-0.93	-0.92	-0.69	0.82	0.85	1.00

Notes: This figure shows the correlation between a state's Jim Crow index, Historical Racial Regime index ([Baker, 2022](#)), number of Jim Crow laws, voter participation ([ICPSR, 1999](#); [Bernini et al., 2023](#)), quality of Black schools ([Card and Krueger, 1992](#)), and causal 1870-ancestor state effects on Black Americans' 1940 years of education as shown in Panel A of Appendix Figure [A.6](#).

C.1.2 New Database on Jim Crow Laws

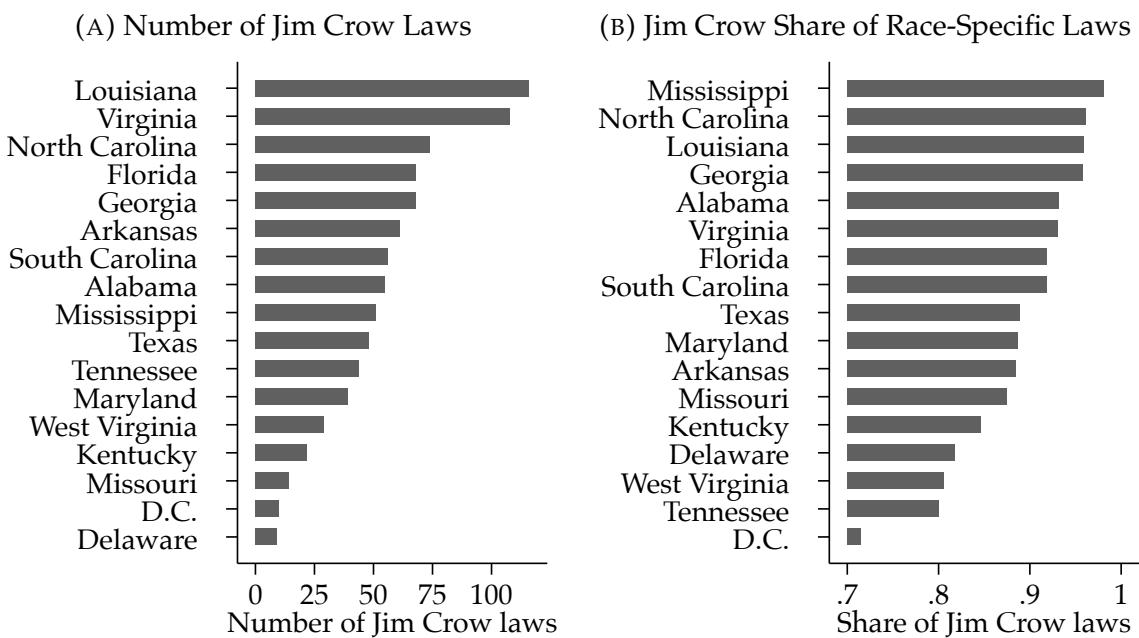
We collect information from 800 Jim Crow laws from four sources, covering both race-specific and “race-blind” Jim Crow laws. We digitize a comprehensive collection of laws that refer to race and color by state in 1950 [Murray \(1950\)](#). We digitize those laws and categorize them as discriminatory, anti-discriminatory, or neutral. We restrict our sample to discriminatory laws and further categorize the domain they pertain to, such as education, suffrage, or employment. Our remaining sources add Jim Crow laws that made no explicit mention of race. We collect laws that limited the geographic mobility of Black Americans and regulated their employment arrangements from [Roback \(1984\)](#) and [Cohen \(1991\)](#). We further collect laws that restricted Black suffrage from [Walton et al. \(2012\)](#). Figure [C.32](#) shows the number of total Jim Crow laws passed by each state until 1950. Figure [C.33](#) shows the distribution over years in which Southern governments passed laws of different types.

FIGURE C.31: Jim Crow Laws by Type



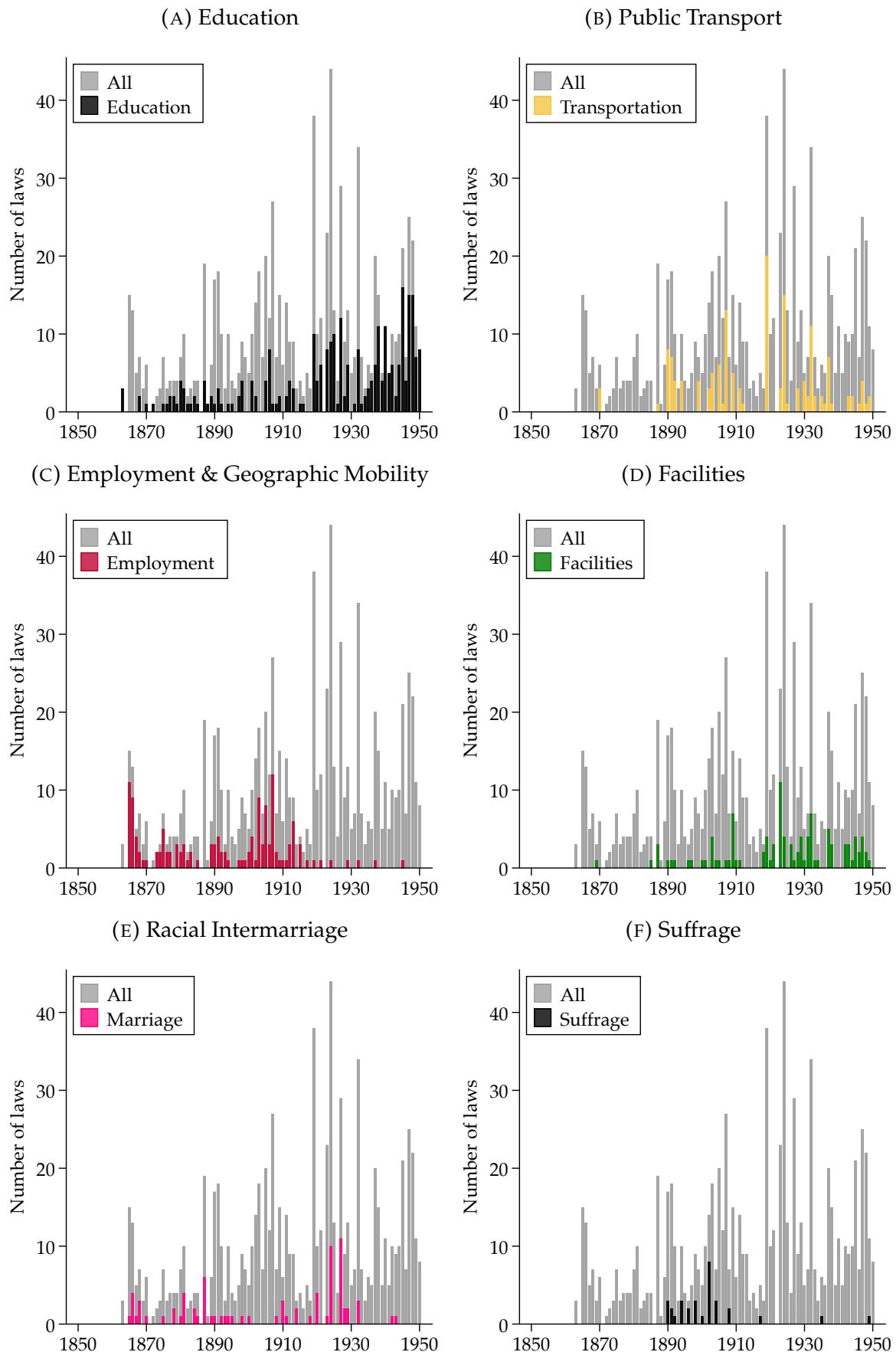
Notes: This figure shows the number of Jim Crow laws across Southern states that pertain to each category. See Data Appendix C for details on the data.

FIGURE C.32: Jim Crow laws by State



Notes: Panel A of this figure shows the cumulative number of Jim Crow laws passed by state until 1950. Panel B shows the anti-Black discriminatory share of all race-specific laws a state passed until 1950.

FIGURE C.33: Annual Jim Crow Laws Passed Across the South by Type



Notes: This figure shows the number of Jim Crow laws passed by type across all Southern states and years.

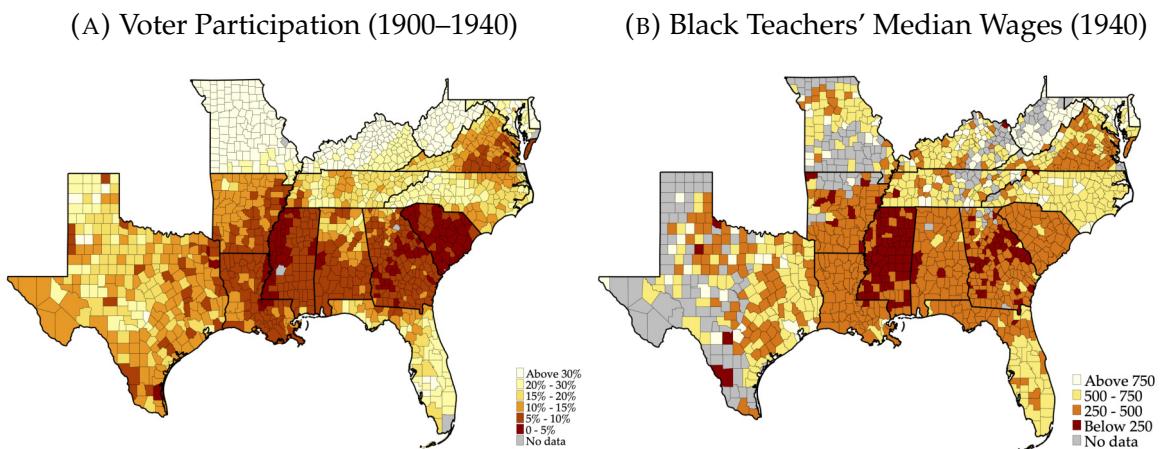
C.1.3 Other Data on Jim Crow Regimes

Historical Racial Regime (HRR) index. As an alternative measure of a state’s Jim Crow intensity, we use the HRR index (Baker, 2022). This index “measures different manifestations of the US racial regime across different historical periods—slavery and Jim Crow—and is based on state-level institutions including slavery, sharecropping, disfranchisement, and segregation.”

Voter participation. As a second alternative measure for the intensity of Jim Crow regimes, we compute a county’s aggregate voter participation in decennial presidential elections in the South from 1900 to 1940 (ICPSR, 1999; Bernini et al., 2023). We divide the total number of votes cast in each election by a county’s total population (see panel A of Figure C.34). Data on the number of votes cast by race are not available. Appendix Figure A.12 shows border discontinuities in voter participation.

Black school quality index. Last, as a third alternative measure for the intensity of Jim Crow regimes, we construct an aggregate measure of Black school quality in the South (Card and Krueger, 1992). We extract a principal component from three measures of Black school quality by state prior to 1940: student-teacher ratios, term lengths, and teacher wages. We also use individual-level data on Black teachers’ wages from the 1940 census to assess whether or not Black school quality differed sharply across state borders (see panel B of Figure C.34). Appendix Figure A.12 shows border discontinuity estimates in teacher wages.

FIGURE C.34: Voter Participation and Black Teachers’ Wages Differ Sharply Across States



Notes: Panel A of this figure shows the average fraction of each county’s population that cast a vote in decennial Presidential elections between 1900 and 1940. Panel B of this figure shows the median annual wage income of Black teachers in the 1940 census for each Southern county. Results for the Black-white ratio in teachers’ median annual wage income are very similar and available upon request. Appendix Figure A.12 shows border discontinuity estimates in both outcomes.

C.2 Census Data

We use the publicly available full-count US 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.

C.3 Census Linking

We use crosswalks for individuals across 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, surname, 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 link not only individuals, but also create family trees based on census information on family interrelationships among people in the same household. Table C.14 shows the linking rates for our family trees as compared to the linking rates among

TABLE C.14: Family Tree’s Linking Rates

	Individual		Family
	Adjacent only	Incl. non-adjacent	
1870 to 1900	12.8%	25.9%	27.6%
1870 to 1910	3.5%	19.4%	24.8%
1870 to 1920	1.1%	12.3%	26.0%
1870 to 1930	0.3%	6.2%	14.2%
1870 to 1940	0.1%	3.1%	9.8%

Notes: This table shows the linking rates for Black men from 1870 to each decade from 1900 to 1940. The first column shows the linking rate when conditioning on finding a person in each adjacent decade (e.g., 1870 to 1900 would require a person to be linked from 1870 to 1880 and from 1880 to 1900). The second column shows the linking rate when allowing for intermediate decades to be skipped (e.g., 1870 to 1900 would require a person to be linked *either* from 1870 to 1880 and from 1880 to 1900 *or* from 1870 to 1900 directly). The third column shows the linking rate when linking either the individual or their ancestors or descendants in the same household (again, allowing intermediate decades to be skipped).

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.

C.4 Identifying Descendants of the Free and Enslaved

C.4.1 Main Method: Linking Historical Census Records

Figure C.35 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.

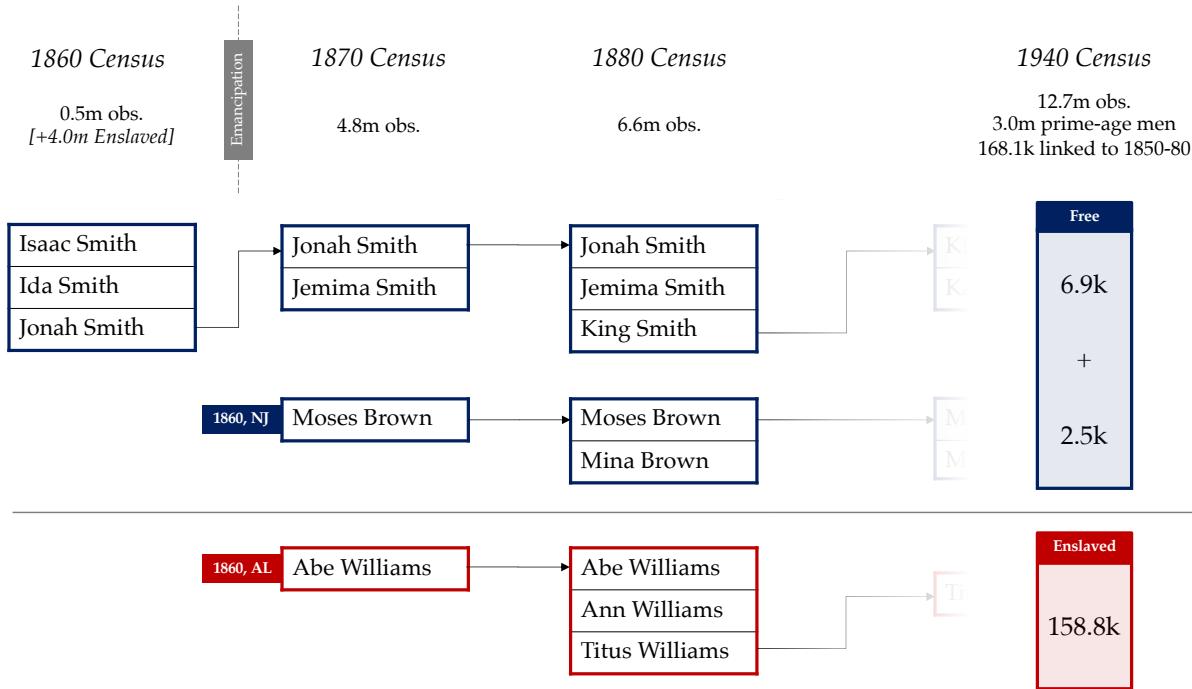
In addition to linking, we use information on place and year of birth in our classification algorithm. All Northern states had begun banning or restricting slavery by 1804—some of them decades earlier. Any Black person born in those states was either free upon birth or would be emancipated by a certain 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 C.16, 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

FIGURE C.35: Illustration of Our Free-Enslaved Classification Algorithm



Notes: This figure illustrates our new method to identify descendants of the Free and 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.

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 C.15 shows that, if anything, the linking procedure biases the Free-Enslaved gap toward zero.

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 due to name changes at marriage. Our classification algorithm identifies descendants of the Free mainly through whether they can be

TABLE C.15: Assessing Linking Bias

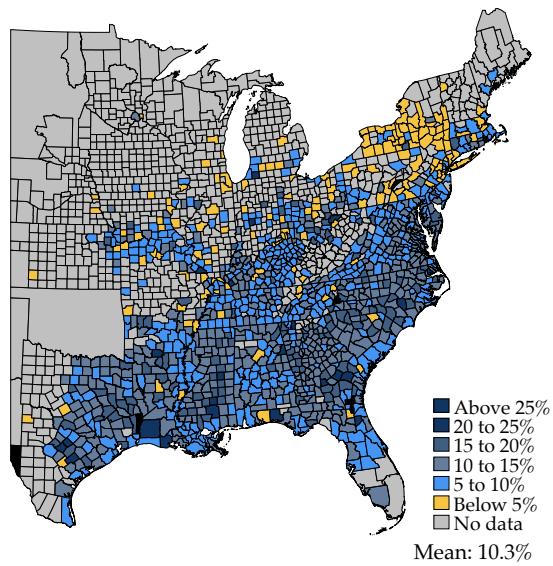
	Free (1860)			Enslaved (1870)		
	Linked	Population	Δ	Linked	Population	Δ
Literacy (%)	65.1	66.8	-3%	20.4	20.4	0%
Occupation Score	6.0	6.1	-1%	3.7	3.8	-1%
Real property (\$)	1,217	1,230	-1%	1,400	1,270	10%
Personal property (\$)	312	316	-1%	312	293	6%
Lives in North (%)	45.1	52.1	-13%	7.8	8.2	-4%
Lives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

Notes: This table shows that there is little selection into the linked sample. If anything, the linked sample is negatively selected for the Free and positively selected for the formerly Enslaved, attenuating the Free-Enslaved gap toward zero. The left panel compares the Free who can be linked to any future decade to the entire 1860 population (which only contains free Black Americans). The right panel compares our linked sample to the 1870 population (89 percent of whom were enslaved until 1865).

linked back to 1850 or 1860, which could lead to an almost mechanically higher socioeconomic status. We addressed this concern in Section 3.4 (see Figure 2).

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 C.36 addresses this concern by showing counties' average likelihood of a resident in 1870 being linkable to the 1940 census. Linking rates are similar across the country except for the most sparsely populated counties in the North (which do not contribute to most of our main results).

FIGURE C.36: Linking Rates by County from 1870 to 1940



Notes: This figure shows the average linking rate for Black prime-age (20–54) men in 1870 to 1940. Only counties with a Black population of at least 50 prime-age men in 1870 are included.

TABLE C.16: Abolition of Slavery in the North

Year	State	<i>De Jure</i> Abolition of Slavery	<i>De Facto</i> Number of Slaves	
			Year	Total
1777	Vermont	Slavery was banned immediately upon founding of Vermont (Constitution of Vermont, 1777).	1790	0 ³⁵
			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
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	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. ³⁵	1790	—
			1800	135 (21%)

TABLE C.16: Abolition of Slavery in the North

Year	State	<i>De Jure</i> Abolition of Slavery	<i>De Facto</i> Number of Slaves	
			Year	Total
	Illinois		1810	429 (28%)
	Michigan		1820	1,106 (40%)
	Wisconsin		1830	788 (5%)
	Minnesota		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 1800 1810 1820 1830 1840 1850	21,324 (82%) 20,343 (66%) 15,017 (37%) 10,088 (26%) 75 (0%) 4 (0%) 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). ³⁶	1790 1800 1810 1820 1830 1840 1850	11,423 (81%) 12,422 (74%) 10,851 (58%) 7,557 (38%) 2,254 (11%) 674 (3%) 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 \(1790–1850\)](#).

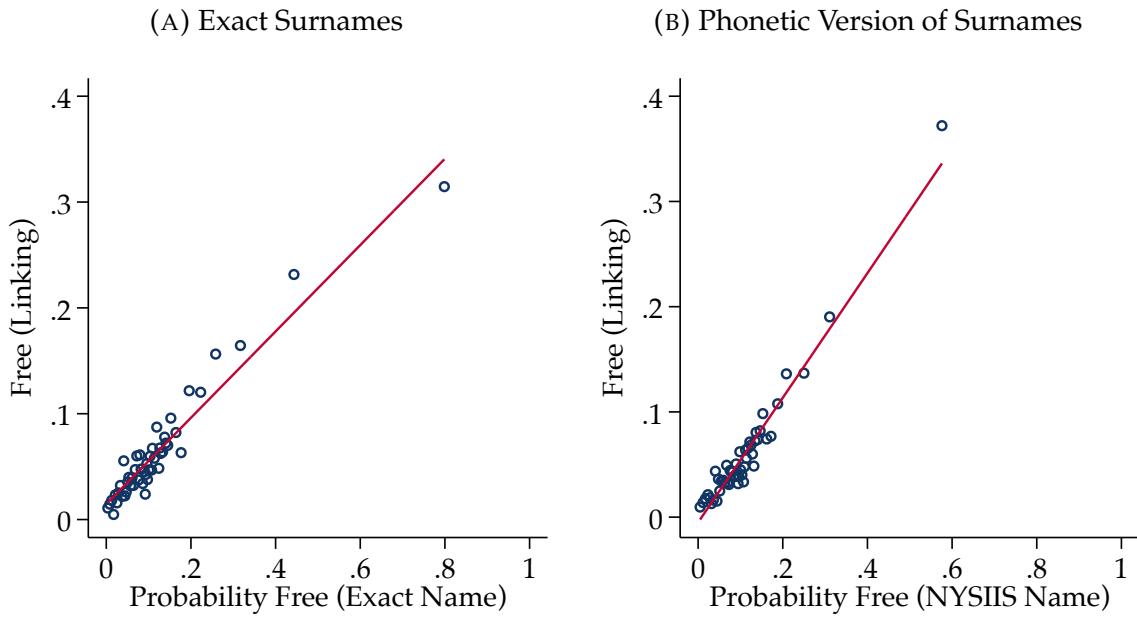
C.4.2 Alternative Method of Free-Enslaved Classification: Distribution of Surnames

While our main method provides a high-accuracy classification of descendants of the Free and 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 Enslaved based on surnames. Figure C.37 shows that the name-based measures are highly correlated with the free status based on our preferred measure, though they are attenuated as expected.

³⁶While the 1790 census states that 16 slaves were in Vermont that year, this is likely an error.

³⁶There is some evidence that after 1804, some Black Americans were sold to slave states before they reached the age to be emancipated ([Armstead et al., 2016](#), p.104).

FIGURE C.37: Comparing Name-Based and Linking-Based Measures



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 surnames.

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

We compute the relative frequency of each surname before and after 1865. We then create a measure of how likely a person is to descend from the Free by dividing their surname's relative frequency before 1865 by its relative frequency after 1865. For example, the surname 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 entered 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 surname 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 person descends from the Enslaved or not. 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 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 surname 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 surname 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

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

where $\#DuBois_t$ is the number of individuals with the surname 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. Table C.17 shows a Black person's probability of descending from ancestors who were enslaved until 1865, given their surname.

To allow for misspellings, we also compute this measure based on the phonetics of surnames. Specifically, we transform surnames using the New York State Identification and Intelligence System (NYSIIS) phonetic code. For example, the surnames "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.

C.5 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

TABLE C.17: Selected Surnames and Enslavement Status

Surname	Likelihood Enslaved
Wanamaker	0%
Du Bois	1%
Cumberland	2%
Dewitt	6%
Radcliffe	10%
McCollins	16%
Dupas	21%
Freemann	28%
Butcher	44%
Freeman	66%
Tubman	70%
Baptiste	85%
Jackson	86%
Broom	87%
Douglass	87%
Johnson	87%
Smith	89%
Carter	90%
Robinson	90%
Hamilton	91%
King	91%
Morrison	91%
Williams	91%
Hughes	92%
Jefferson	92%
Marshall	92%
Baldwin	94%
Jordan	94%
Lincoln	95%
Knowles	96%
Washington	96%
Cooks	97%
Broadnax	99%
Boykins	100%
Doyley	100%
Gadson	100%
Freedman	100%
Merriweather	100%
Rockingham	100%

Notes: This table shows estimates of the probability of descending from enslaved Black Americans by surname (conditional on being Black). Some examples are taken from [Clark \(2014\)](#), who lists a number of surnames that “sound classically English” but tend to be predominantly Black today, suggesting that they were likely “adopted in the slavery era from master whose own families died out or left few descendants.” Consistent with that idea, our estimates suggest that Black people with those surnames are almost certain to descend from ancestors who were enslaved until the Civil War.

families who migrated to the US 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 C.4.

C.6 Credit Bureau Sample

We analyze data from a major US credit bureau, which includes comprehensive monthly credit reports for individuals from January 2010 to present. These reports, updated on the final Tuesday of each month, contain information from various sources, such as financial institutions, debt collection agencies, and public records, along with proprietary data. Our focus is on the March 2023 snapshot, which aligns with our findings from a comparable pre-COVID sample.

Our sample is restricted to Black prime-age (20-54) men. The credit bureau uses a predictive method to determine race, based on a person's name and their neighborhood (nine-digit ZIP code). This method, given the detailed geographic information it leverages, is far more accurate than common proxies that rely solely on surnames. Using a separate dataset—our Social Security mortality records—we find that surnames capture 22 percent of the variation in whether a person is Black or not; nine-digit ZIP codes capture 76 percent; and both combined capture 90 percent.

The bureau combined our probabilistic surname-based classification of Free-Enslaved status of Black individuals with their credit reports, subsequently anonymizing the data. We access these anonymized individual-level credit reports for around 550,000 Black prime-age men via a secure server, allowing real-time estimation of the Free-Enslaved gap in employment and credit.

C.7 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](#)). We use additional individual-level data from a major US credit bureau to extend our results to 2023.

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 in-

come 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.
- **Total income, 2019–2023 (credit bureau).** Measures a household’s gross total compensation for the most recent year reported. This measure is estimated based on a proprietary data and prediction models.
- **Disposable income, 2019–2023 (credit bureau).** Measures a household’s income available to spend, invest, or save after accounting for fixed expenses. This measure is estimated based on a proprietary data and prediction models.
- **Hourly job, 2019–2023 (credit bureau).** Measures whether a person is employed as an hourly or salary worker.

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.
- **Credit score, 2019–2023 (credit bureau).** The VantageScore® 3.0 measures a person’s credit health. The score takes into account a rich set of indicators on a person’s financial situation. It ranges from 300 to 850. Scores above 700 are typically considered “good” and scores below 550 “very poor.”

C.8 Neighborhood-Level Outcome Variables

While we cannot link our data to censuses after 1940, we can link the 1940 census to administrative mortality records from 1988 and 2005 using the CenSoc-Numident file ([Goldstein et al., 2021](#)). Importantly, the mortality 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 [C.10](#) 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.

One potential concern with this data may be that many people live in retirement homes, possibly making the neighborhood a less precise proxy of a person's socioeconomic status. To assess this potential issue, we compare the density of deaths with a ZIP code's population density and find that the two are highly correlated ($\rho = 0.91$). We show that our results are robust to dropping ZIP codes that have far higher rates of deaths than predicted by their population density (see Table [B.9](#)).

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.

Demographics

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

C.9 County Characteristics

We compile a dataset on county characteristics combining data from the IPUMS National Historical Geographic Information System (NHGIS, [Manson et al., 2021](#)) and various other sources.

- **Black, 1860 (NHGIS).** Measures the percentage of a county's 1860 population that is Black.
- **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.
- **Farm, 1870 (NHGIS).** Measures the fraction of a county's population living on a farm in 1870.
- **Free, 1860 (NHGIS).** Measures the percentage of a county's 1860 Black population that is free.
- **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.
- **Lynchings, 1883–1941 (Seguin and Rigby, 2019).** Measures the number of lynchings that occurred in a county between 1883 and 1941.
- **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.
- **Population density, 1870 (NHGIS).** Measures a county's 1870 population per square kilometer area.

- **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.
- **School, 1870 (NHGIS)**. Measures the fraction of a county's Black children (ages 6–16) attending school in 1870.
- **Slaves per slaveholder, 1860 (NHGIS)**. The average number of enslaved people per slaveholder.
- **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.
- **Voter participation, 1860–1940 (ICPSR, 1999; Bernini et al., 2023)**. Number of votes cast across parties in each presidential election.

C.10 Nine-Digit ZIP to Census 2000 Crosswalks

The administrative mortality 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.³⁷ 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 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

³⁷Extensive documentation is available online for each TIGER/Line version.

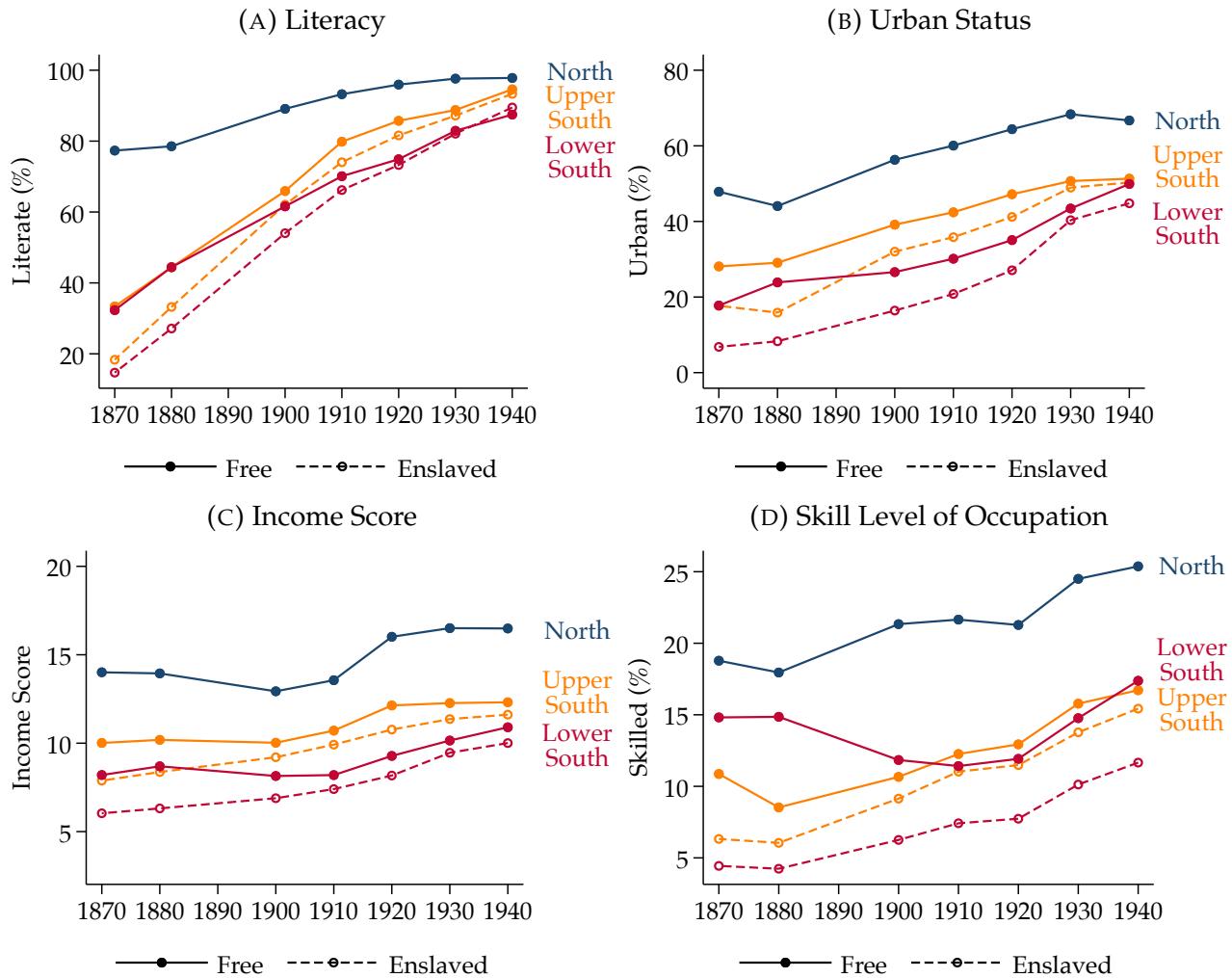
matches were unique, respectively. In cases where a ZIP9 occurs in more than one statistical area, we assign the area that has the largest number of matches with the relevant ZIP9. This yields a one-to-one mapping of ZIP9s to blocks for each TIGER/Line year between 2000 and 2006. However, not all ZIP9s in the Censoc-Numident mortality 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).

C.11 Descriptive Statistics

C.11.1 Socioeconomic Status of Descendants of the Free and Enslaved

FIGURE C.38: 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 Enslaved by their ancestor's region (family's residence pre-1880). Incomes Score uses the Lido score developed by [Saavedra and Twinam \(2020\)](#). In the 1940 census, instead of literacy, we observe the highest year of school or degree completed. We classify individuals who have completed at least two grades of school as literate; others we classify as illiterate. We assign "skilled" to occupations classified as "medium skilled workers" or above by the HISCLASS scheme ([Leeuwen and Maas, 2011](#)); and "unskilled" to others. See Data Appendix C for details on the sample and data.

C.11.2 Migration among Descendants of the Free and Enslaved

FIGURE C.39: County Population of Enslaved and Free (1790)

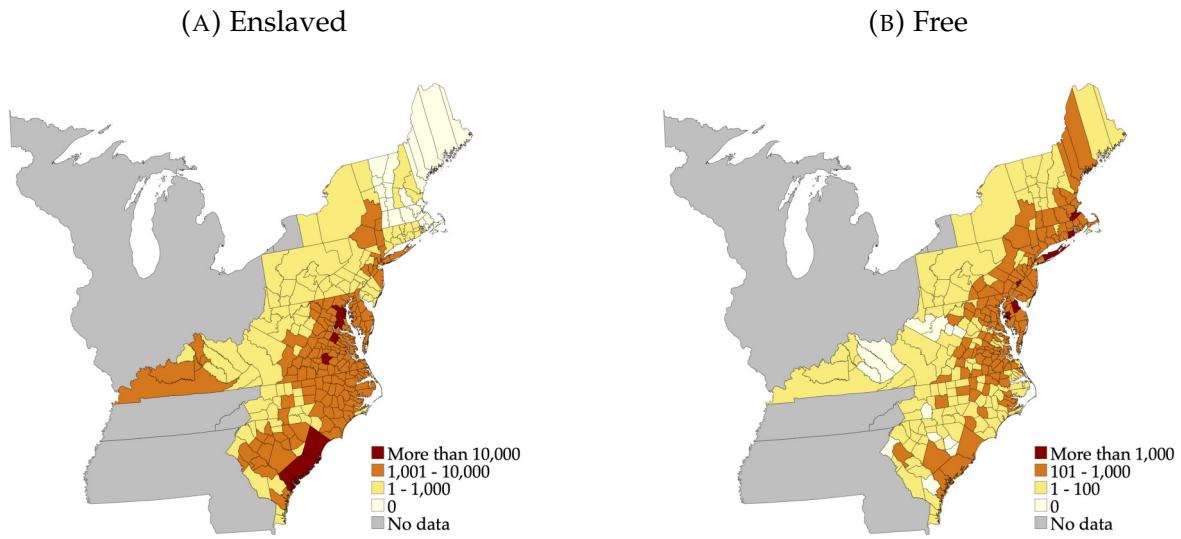
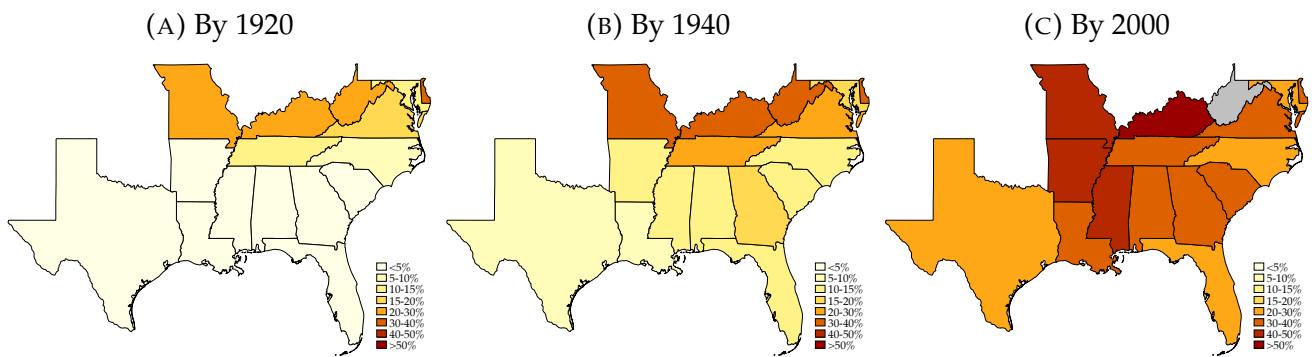
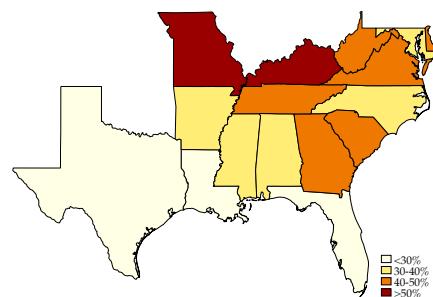


FIGURE C.40: Black Families Leaving the Slave States by 1870 State of Origin



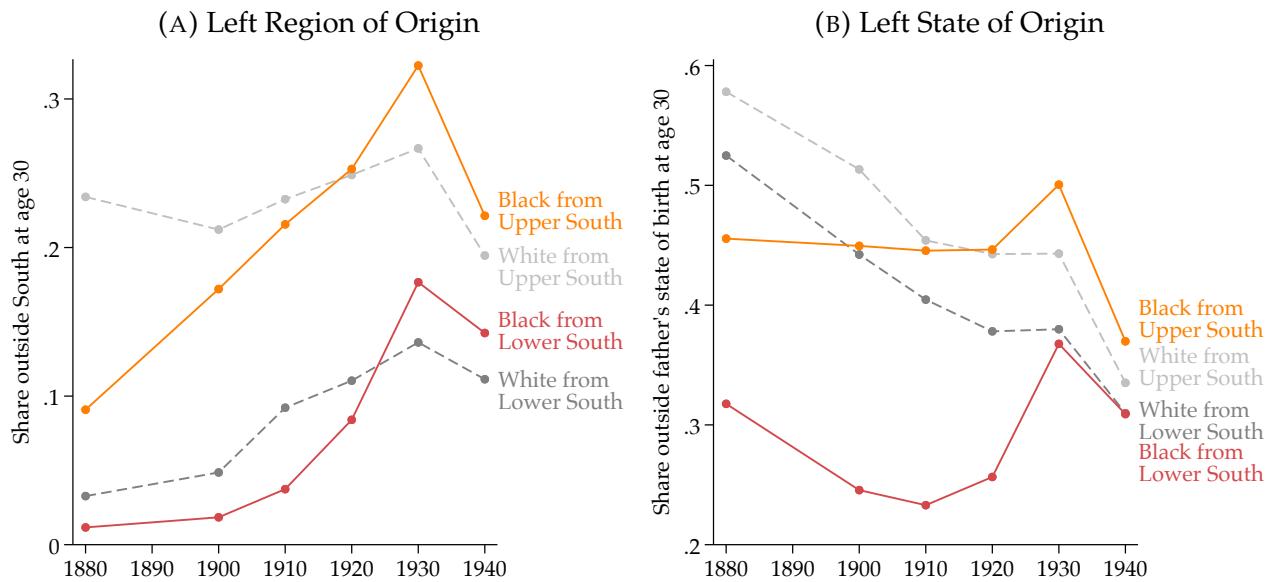
Notes: This figure shows the cumulative fraction of Black families who live outside the slave states, by the state their 1870 ancestor was born. The figure highlights that the first wave of the Great Migration from 1910 to 1940 was mainly an Upper Southern phenomenon (see Panels A and B). Black families with roots to the Lower South only caught up with those rates of migration to the North after 1940 (see Panel C).

FIGURE C.41: Black Families Leaving their 1870 State of Origin by 1940



Notes: This figure shows the fraction of Black families who in 1940 live outside the state in which their ancestors were enslaved. As the state of enslavement, we use the state of birth of formerly enslaved ancestors in the 1870 census.

FIGURE C.42: Long-Term Migration Rates across Regions and States by Race



Notes: This figure shows the fraction of Black and white individuals aged 30 who have migrated from their father's birth region (Panel A) or father's birth state (Panel B) in each census year. The data is derived from the 1850–1940 censuses, focusing on the Southern-born fathers' states of birth, and does not require census linking.

REFERENCES

- (1956): "The Southern Manifesto," *Congressional Record*, 102, 4459–4460, 84th Congress Second Session.
- 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.
- BAKER, R. S. (2022): "The Historical Racial Regime and Racial Inequality in Poverty in the American South," *American Journal of Sociology*, 127.
- BERLIN, I. (1974): *Slaves without masters: The free Negro in the antebellum South*, Oxford University Press.

- BERNINI, A., G. FACCHINI, AND C. TESTA (2023): "Race, Representation, and Local Governments in the US South: The Effect of the Voting Rights Act," *Journal of Political Economy*, 131, 994–1056.
- CARD, D., C. DOMNISORU, AND L. TAYLOR (2022): "The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility," *Journal of Labor Economics*, 40, 39–95.
- CARD, D. AND A. B. KRUEGER (1992): "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics*, 107, 151–200.
- CASCIO, E. U. AND E. G. LEWIS (2022): "Legal Activism, State Policy, and Racial Inequality in Teacher Salaries and Educational Attainment in the Mid-Century American South," Working Paper 30631, National Bureau of Economic Research.
- 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.
- CLARK, G. (2014): *The Son Also Rises: Surnames and the History of Social Mobility*, Princeton University Press.
- COHEN, W. (1991): *At Freedom's Edge: Black Mobility and the Southern White Quest for Racial Control, 1861–1915*.
- 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.
- 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/>.

ICPSR (1999): “United States Historical Election Returns, 1824-1968,” Inter-university Consortium for Political and Social Research.

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>.

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

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/>.

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>.

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.

WALTON, H., S. C. PUCKETT, AND D. R. DESKINS (2012): *The African American Electorate: A Statistical History*, Washington, D.C.: CQ Press.

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

D. MODEL APPENDIX

We think of the effect of being enslaved until 1865 as the expected difference between the descendants of the Enslaved and descendants of the Free holding constant any confounding factors (i.e., “ability”). That is, we define the average “treatment” effect (ATE) at $t = 1$ as

$$\begin{aligned} ATE &\equiv \int (\mathbb{E}[y_{i,1} | s_i = 1, \alpha_{i,0}] - \mathbb{E}[y_{i,1} | s_i = 0, \alpha_{i,0}]) dF(\alpha_{i,0}) = \\ &= \int \mathbb{E} \left[\rho \left(\gamma_{\ell(i,0)}^0 - \delta \right) + \gamma_{\ell(i,1)}^1 \middle| s_i = 1, \alpha_{i,0} \right] dF(\alpha_{i,0}) - \\ &\quad \int \mathbb{E} \left[\rho \gamma_{\ell(i,0)}^0 + \gamma_{\ell(i,1)}^1 \middle| s_i = 0, \alpha_{i,0} \right] dF(\alpha_{i,0}). \end{aligned} \quad (15)$$

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 (2), (3), and (15), the observed Free-Enslaved gap is equal to

$$\mathbb{E}[y_{i,1} | s_i = 1] - \mathbb{E}[y_{i,1} | s_i = 0] = ATE - B, \quad (16)$$

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:

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

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. *A priori*, this assumption 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 freed through general emancipation in the North. Within the remaining half, it is hard to estimate the share of people 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 attained freedom through escape.³⁸ 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 [C.16](#)), 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

³⁸Pennsylvania Abolition Society and Society of Friends Manuscript Census Schedules, 1838. Available in machine-readable form through <https://doi.org/10.3886/ICPSR03805.v1>.

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 D.18 summarizes the discussion.

TABLE D.18: 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.

REFERENCES

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