The Social Construction of Race after Emancipation: US Census Racial Assignment Based on Skin Tone and Socioeconomic Status

Anjali Adukia
University of Chicago

Richard Hornbeck
University of Chicago

Daniel Keniston
Louisiana State University

Benjamin Lualdi
University of Chicago

January 2025

Abstract

We examine the social construction of race during the United States' Reconstruction Era, a critical juncture between slavery and Jim Crow segregation. We show that people with the same detailed skin tone, recorded by the Freedman's Bank (1865-1874), were more likely racialized as White or Mulatto by the 1870 Census if they were wealthier or literate. Our estimates highlight when race was being constructed – or rather, reconstructed – in a period of unfulfilled potential for changing social structure after emancipation, setting a path for racial segregation and continued racial stratification. This endogenous origin of race also has implications for analyses that compare individuals by race or include race as a control variable.

For helpful comments and suggestions, we thank many colleagues and seminar participants at DePaul, LSE, RIDGE-LACEA, UChicago, and Warwick. We are grateful to our data teams at Digital Divide Data and the University of Chicago, as well as to Eric Chyn, John Clegg, Kareem Haggag, and Bryan Stuart for sharing data. This research was funded in part by the Neubauer Family Faculty Fellowship.

The United States' social structure was in flux after slavery and emancipation. The Reconstruction Era (1865–1877) saw efforts to integrate formerly-enslaved people into broader social, economic, and political institutions, until backlash gave rise to state-supported segregation in the Jim Crow Era (1877-1964). Without the legal distinction between "free" and "enslaved," segregation relied on constructed racial groupings. People had to determine who was what race – and what race was. This social construction of race, particularly during the critical period after emancipation, set the stage for racial segregation and continued racial stratification.

In this paper, we show that people with the same physical skin tone were racialized differently in the 1870 Census based on their wealth and other proxies for socioeconomic status (SES). Census enumerators visited households in 1870, categorizing people as "Black," "Mulatto," and "White."[1] Census instructions noted: "Be particularly careful in reporting the class Mulatto" to include "quadroons, octoroons, and all persons having any perceptible trace of African blood" on which "important scientific results depend" (U.S. Census Office, 1870, p. 10). The 1870 Census reflects enumerators' judgment along with self-reported information.[2]

The Freedman's Savings Bank (1865-1874) recorded depositor "complexion" to help identify customers when they returned to the bank, which was generally their perceived skin tone (e.g., dark, dark brown, brown, light brown, light). Recorded data sometimes coincided with racial categories (e.g., black, colored, nearly black, mulatto, nearly white, white) and occasionally included other physical descriptives (e.g., distinctive scars or pox marks).[3]

Table 1 provides a motivating tabulation of the data, comparing depositors' complexion and their race in the 1870 Census. Darker skin tones are associated with being racialized as Black (along the diagonal in columns 2 – 4), but there is substantial variation in skin tone within Census race (looking down the columns) and substantial variation in Census race within skin tone (looking across the rows).[4] We examine correlations with skin tone

[1]The two other racial categories in 1870 were "Chinese" and "Indian" (Native American), which are very rare among Freedman Bank depositors and we exclude from our analysis. The 1870 Census labels this variable "Color," but the Census classifications are racially focused in practice and changed over time with evolving racial structures (Hochschild and Powell, 2008; Humes and Hogan, 2009).

[2]For example, in discussing the recording of occupation, the 1870 Census instructions note: "You are under no obligation to give any man's occupation just as he expresses it. If he cannot tell intelligible what he *is*, find out what he *does*, and characterize his profession accordingly [emphasis original]" (U.S. Census Office, 1870, p. 14).

[3]Union Army records on "complexion," by contrast, predominately refer to African American soldiers as "Black." We thank John Clegg for providing tabulations from the African American Civil War Soldiers Project.

[4]Note that "Yellow" was a descriptor at the time for a light-skinned African American person. The Freedman branch in New York City attracted immigrant depositors, but Freedman branches were generally across the South with less immigration. Almost all immigrants are white in the complexion data and in the Census (e.g., people born in Italy were 99.7% "White" in the full 1870 Census), such that our estimates are not sensitive to excluding the Freedman Branch in New York City or all foreign-born people in the Census.

within Census race, but emphasize the selection bias in such correlations that arises from endogenous Census racial assignment. Our main focus is on the variation across rows: for people in the same complexion group, the systematic variation in who was racialized as Black, Mulatto, or White.

Racial assignment in the Census provides a view into how people were being racialized at the time. Racial perceptions, and the resulting racial categorization, substantially impacted how people were treated by the law and society more generally. The Census itself, as an instrument of the government, also had a role in supporting the ongoing construction of race.

We find evidence for social construction of race: among people with the same skin tone, those with higher wealth or literacy are more likely to be racialized as Mulatto or White. We estimate selected racialization along each of three margins: White or Mulatto vs. Black, White vs. Mulatto or Black, and Mulatto vs. Black. The effects of wealth and literacy on racialization are largest for people with light skin tones. We also estimate selected racialization based on occupation, though a particular advantage of the 1870 Census is its reported data on physical wealth that is central to a family's socioeconomic status and security (along with the opportunities associated with being literate).

We compare depositors within Freedman Bank branch (Figure 1), which are mostly in larger cities in the South at the time, and explore geographic heterogeneity in the effects across counties (e.g., with variation in racialized violence). We see similar effects by gender and for adults and children. Skin tone varies within family, among families with multiple Bank depositors, but assigned Census race varies less often within family, which reflects social construction also based on family relationships. Consistent with broader social influences, we also report differential racialization based on naming patterns.

We devote a great deal of attention to record linkage, which has the potential to bias our estimates. We hand-link Freedman Bank records to the contemporary 1870 Census using biographical information for the depositor (name, age, birth state), along with other household member names that give us greater confidence in links, and iterate on these hand-links with a machine-learning model. Our estimates are not sensitive to restricting the sample to links with increasingly supportive supplemental information.

We also bound the potential bias from unobserved physical characteristics, which might be correlated with both racial assignment and socioeconomic status. This bounding exercise leverages the observed influence of our detailed skin tone data, compared to less-detailed skin tone categories (following Altonji, Elder and Taber, 2005; Cinelli and Hazlett, 2020). The estimated bounds are close to our baseline estimates, reinforcing the important role of wealth and other proxies for SES in how people were racialized. Our observed proxies for SES can still be correlated with unobserved social cues, such as how people dress or speak,

which have the same interpretation.

The primary implication of these estimates is to illustrate the social construction of race during this critical period after emancipation, separating people along socioeconomic lines, which set a path for racial segregation after emancipation and the collapse of Reconstruction. Following early successes during Reconstruction, racial backlash led to its collapse (Logan, 2020; Chyn, Haggag and Stuart, 2024) and racism adapted (Engerman, 2003). After Reconstruction, an increasingly dichotomous racial categorization not only supported white social control in the absence of enslavement (Washington, 2011), but was also reinforced through black collectivism to counter segregation (Reece, 2021). Racial progress stalled as segregation became entrenched through the Jim Crow era, limiting economic mobility and convergence in racial wealth gaps even after the Jim Crow era ended (Chyn, Haggag and Stuart, 2022; Althoff and Reichardt, 2024; Derenoncourt et al., 2024). The systemic nature of continued racial divisions and disparities are highlighted in stratification economics (Darity Jr., 2005; Darity Jr., Mason and Stewart, 2006; Darity Jr., Hamilton and Stewart, 2015; Chelwa, Hamilton and Stewart, 2022).

A secondary implication of race being endogenous is to complicate the interpretation of estimated differences within race. We use a selection model to simulate selection bias when "controlling for race," and show corresponding empirical patterns. In our data, darker skin tones are associated with lower wealth (and other proxies for lower socioeconomic status). Within Census race, however, there is a substantially attenuated relationship between skin tone and wealth. One reason is that among Black people (i.e., those racialized as Black in the 1870 Census), people with light brown skin are often poor because if they were wealthier then they would often not be considered Black. Similarly, among White people (i.e., those racialized as White in the 1870 Census), people with light brown skin are often wealthy because if they were poorer then they would often not be considered White. By contrast, non-classical measurement error (e.g., if the Freedman's Bank recorded lighter skin tones for people when they appeared wealthier) would bias estimates in the opposite direction from our results, creating a stronger relationship between measured skin tone and Census wealth and attenuating our estimated effect of wealth on racial assignment within skin tone. We also show little influence from classical measurement error in skin tone, using a sub-sample of depositors with two measures of skin tone from opening accounts twice.

Our main contribution is to show the influence of socioeconomic status on racialization during this critical juncture after emancipation, leveraging uniquely-detailed historical data on physical skin tone from the Freedman's Bank that we link to the 1870 Census. This paper joins a large literature on the social construction of race and its consequences across Anthropology, Critical Race Theory, Economics, Health, History, Law, Political Science,

Psychology, and Sociology.[5] Social cues influence the racial perception of faces in modern lab settings (Freeman et al., 2011), and how people are racialized can vary with changing life circumstances (Penner and Saperstein, 2008; Saperstein and Penner, 2012, 2016).

Our core question is cross-sectional – how people are racialized, and how these racial categories were defined during Reconstruction – which is feasible to explore because of our uniquely-detailed data from the Freedman's Bank. Without data on skin tone, a complementary literature explores changes in how people are racialized. Historical analyses of Census-to-Census links examine changes in linked individuals' race from Black to Mulatto (Saperstein and Gullickson, 2013), from Black to White (Dahis, Nix and Qian, 2020), and from Mulatto to White (Mill and Stein, 2016), for which some emphasize race affecting occupation and others emphasize occupation affecting race. Potential errors in Census-to-Census links, particularly when recorded race is different, can generate spurious results. We hand-link contemporaneous records from the 1870 Census and Freedman's Bank, using detailed supplemental information in the contemporaneous Bank records. Contemporaneous links are also useful because someone's physical appearance can change over time. Our analysis of racialization in the 1870 Census, the only Census during the Reconstruction Era, also captures this historically critical period between emancipation and Jim Crow.

Differences in skin tone have long been associated with differing societal treatment, under enslavement in the United States (Keith and Herring, 1991; Bodenhorn, 2015; Reece, 2018) through the present day (Wade, Romano and Blue, 2004; Burch, 2015).[6] McGee (2025) develops a theoretical model for how elites use skin color in the construction of race to establish and maintain their socioeconomic position, which is consistent with differences by Census race in the US and differences by skin tone in Brazil within race. A long literature explores the influence of skin tone in other countries, and Latin America in particular, where racial categories are often more fluid than in the United States (Lovell and Wood, 1998; Villarreal, 2010; Telles and Paschel, 2014; Rangel, 2015; Monk Jr., 2016; Davenport, 2020; Laudares and Valencia Caicedo, 2023; Woo-Mora, 2024). In contrast to literature in the US

[5]See, for example, Boas (1912); Montagu (1945); Rogers (1952); Hahn, Mulinare and Teutsch (1992); Haney López (1994); Witzig (1996); Bonilla-Silva (1997); Zuberi (2001); Darity Jr. (2005); Smedley and Smedley (2005); Goldsmith, Hamilton and Darity Jr. (2007); Penner and Saperstein (2008); Davis (2010); Charles and Guryan (2011); Freeman et al. (2011); Morning (2011); Saperstein and Penner (2012); Saperstein and Gullickson (2013); Penner and Saperstein (2013); Omi and Winant (2014); Darity Jr., Hamilton and Stewart (2015); Hersch (2018); Reece (2019); Dahis, Nix and Qian (2020); Davenport (2020); Ioannidis, Powe and Yancy (2021); Chelwa, Hamilton and Stewart (2022); Francis, Hardy and Jones (2022); Abramitzky et al. (2023); Rose (2023).

[6]Abramitzky et al. (2023) highlight different outcomes for children classified as Mulatto or Black, within the same household, over the late-19th and early 20th centuries. Noghanibehambari and Fletcher (2025) find that children classified as Black, who appear in linked samples as White adults in the 1940 Census, go on to live longer.

and elsewhere that explores differences in socioeconomic status by skin tone, controlling for race, our analysis shows that race is an outcome – influenced by socioeconomic status among people with the same physical skin tone.

Colorism refers to discrimination based on skin tone, which typically manifests in people with lighter skin receiving preferential treatment.[7] Research often estimates disparities for people with darker skin, within race, which we show substantially understates the true effects of skin tone when race itself is endogenous. For example, estimates that wages are higher among lighter-skinned African Americans, as in Goldsmith, Hamilton and Darity Jr. (2006, 2007); Kreisman and Rangel (2015), would *understate* the effect of skin tone when being classified as African American is endogenous to income.[8]

Despite some recognition that "race is a social construct," general statistical practice is still to "control for race" along with other characteristics (Sen and Wasow, 2016; Rose, 2023). We emphasize that examining differences in outcomes, within racial groups, is complicated to interpret when racial groups are themselves an outcome.[9]

Our takeaway is not just that "race is a social construct," but that we can now for the first time see the historical construction of race along SES lines in this critical period – immediately following emancipation – that set the stage for racial segregation and continued racial stratification during the Jim Crow Era and after. Using unusually detailed skin tone data from the Freedman's Bank, we show the socioeconomic influence on mappings between skin tone and racial categories (Black, Mulatto, White). Our estimates highlight the fluidity of socially constructed racial classifications, challenging essentialist views of racial identity, which have had enduring impacts on academic literature and social structure.

[7]Colorism has been documented in many settings, both within and across racial groups, including within families and ethnically homogeneous groups without separate racial groupings (Keith and Herring, 1991; Landor et al., 2013; Dixon and Telles, 2017). Skin color predicts life outcomes including occupation, educational attainment, criminal legal system involvement, mental health, and marital outcomes (Johnson, 1934; Ransford, 1970; Harburg et al., 1978; Hughes and Hertel, 1990; Seltzer and Smith, 1991; Johnson, Bienenstock and Stoloff, 1995; Krieger, Sidney and Coakley, 1998; Klonoff and Landrine, 2000; Bowman, Muhammad and Ifatunji, 2004; Majumdar, 2004; Goldsmith, Hamilton and Darity Jr., 2006; Hersch, 2006; Gyimah-Brempong and Price, 2006; Monk Jr., 2014; Burch, 2015; Diette et al., 2015; Hunter, 2016; Blake et al., 2017; Dixon and Telles, 2017; Hersch, 2018; Louie and Wilkes, 2018; Uzogara, 2019).

[8]Racial mis-classification in administrative data, particularly to avoid detection of discrimination, can also bias estimates of disparate treatment by race (Luh, 2023; Finlay, Luh and Mueller-Smith, 2024; Baron et al., 2024).

[9]Among Native Americans and immigrants to the United States, selective ethnic identification can also create biases (Alba and Chamlin, 1983; Lieberson and Waters, 1993; Waters, 1994; Alba and Islam, 2009; Duncan and Trejo, 2011, 2017; Duncan et al., 2020; Kosack and Ward, 2020; Eschbach, Supple and Snipp, 1998). Census enumerators in 1930 vary systematically in their assignment of people with Mexican ancestry to be "Mexican" or "White" (Duncan and Trejo, 2023), though data on skin tone are unavailable to implement our analysis. There can also be selective exit from ethnic groups over generations if more-educated people marry lighter-skinned people (as in Brazil Schwartzman, 2007), with physical changes over generations, whereas we focus on differential racialization for people with the same skin tone.

I Data Construction and Summary Statistics

I.A Skin Tone Data

We use data on skin tone from the Freedman's Savings Bank (1865-1874). The Freedman's Bank was chartered by Congress in 1865 and marketed to newly-emancipated people, though others also used the Bank. The Bank accommodated small deposits and paid interest. The Bank was not formally connected to the Freedman's Bureau and was not financially backed by the US government, though some promotional materials fostered confusion (Célérier and Tak, 2024). The Bank was allowed to open many branches, unusual for the time, and was initially restricted to investments in US treasuries before successful lobbying by the White leadership of the Bank to invest in riskier assets. After many of these loans failed, in railroads and ventures of the White leadership's associates, the Bank collapsed in 1874 and depositors eventually received at most 62% of their final balance. W.E.B. DuBois (1903) describes the Bank's failure as setting back "the thrift of the freedmen" as much as "ten additional years of slavery could have" (DuBois, 1903, p. 37). The Bank was described by Frederick Douglass as "the black man's cow and the white man's milk" (1874), and Congress declined to compensate depositors' losses.

When someone opened an account with the Bank, the branch cashier filled out a depositor record (Appendix Figure 1). This included the depositor's name, birth state, age, and names of household members along with other information. The form included a field for "complexion," which Osthaus (1976, p. 82) discusses as helping to identify the person if they lost their deposit book. There are 368 unique string entries for complexion in our sample, reflecting spelling variations and less common physical descriptives, which we standardize into 27 detailed skin tone categories (reported in Table 1). Some coincide with racial categories, which we exclude from some specifications and show that use of racial categories attenuates our baseline estimates.[10] Skin tone was the main identifying physical characteristic recorded by the Bank, with only occasional references to other characteristics (e.g., distinctive scars or pox marks). The Bank's choice of skin tone as a primary identifying characteristic is consistent with a large literature on the use of skin color from young ages to distinguish between individuals (Kessen and Weiskopf, 1976; Katz and Kofkin, 1997; Hirschfeld, 2008; Kinzler and Spelke, 2011).

Our main specifications compare people within the 27 detailed skin tone categories listed in Table 1, which are many more categories that are typically available to researchers, but

[10]We also find that Freedman Bank branches with Black cashiers, listed by Osthaus (1976), record more detailed descriptions of skin tone rather than broader or more racial categories. There is insufficient information on Census enumerators to find their recorded race in the 1870 Census, though we expect the Census enumerators were mostly White (as US Marshals and Assistant Marshals in 1870).

we also use a "skin tone index" in some analysis. To create this index, where higher values generally reflect darker skin tones, we use our links to the 1870 Census. For each skin tone category, we define the index as the share of people with that skin tone that are Black in the Census (times one) plus the share of people with that skin tone that are Mulatto in the Census (times one-half). Any such index is a somewhat arbitrary scaling, but provides a continuous measure across categories that is useful for parts of our analysis and reported in column 5 of Table 1.[11] But our main analysis focuses on the variation across rows: for people with a particular skin tone, who is racialized as Black or Mulatto based on their other characteristics.

We observe two entries for complexion data for 650 depositors in our sample, when the depositor opened multiple accounts. For 66% of these depositors, the detailed skin tone category is the same from both records. For an additional 14% of these depositors, the skin tone index values are within 0.05. Figure 2 plots the two skin tone index values against each other, where the size of the circle is proportional to the number of depositors. We use these multiple measurements to instrument for one with the other in assessing the impacts of measurement error in skin tone.

I.B Linkage to 1870 Census Records

We hand-link the Bank depositor records to the 1870 Census, iterating with a machine-learning model to reduce errors and expand the sample. We start with 44,077 linkable depositor records that have the depositor's age and birth state.

We first generate candidate links in the 1870 Census that have the same birth state and gender, age within +/- 10 years, and a Jaro-Winkler string distance between first and last names ≤ 0.3.[12] These initial candidate links allow for fairly wide differences because of the high proportion of newly-emancipated people in our sample that have lower levels of numeracy and literacy. The additional supporting information can suggest promising links for records that are less close matches on age and depositor name. We keep at most the top 20 candidate links.[13]

We then send these candidate links to our data team, where two people independently

[11] When race is endogenous to SES, the ordering of skin tone categories will not be affected if skin tone is also negatively correlated with SES (as we find in the data).

[12] We include exact matches on first initial, assigning a Jaro-Winkler distance of 0.3. For age, we adjust for when the depositor account was opened: if someone reports being 33 years old when they opened a bank account in 1873, we consider candidate people that are between ages 20 and 40 in the 1870 Census. We assign gender for the depositor records based on first name. For immigrants, predominately in the New York City branch, we use their birth country rather than birth state.

[13] We sort the candidate links as follows: by last-name string distance, then first-name string distance, spouse string distance, child string distance, and the difference in age between the depositor record and census record.

hand-link the records. People can link the depositor record to a unique Census record (1:1), or link a depositor record to multiple potential Census records (1:m), or make no link to the Census. For cases where the two people make different links, we send these to a third person for reconciliation.

Our hand-linking process takes advantage of additional information beyond what is typically used in more algorithmic linkage across Censuses (Bailey et al., 2020; Abramitzky et al., 2021). Especially useful for our contemporaneous links are the names of other household members: for adults, their spouse and children; for children, their parents and siblings. We also show our data team members the Census race, combining Black and Mulatto into a "black/non-white" category, along with a coarse grouping of depositor complexion (black, ambiguous, white).[14] Data team members may be more likely to consider Black and White links for those in the ambiguous category, with lighter skin tones, but not differentially with SES characteristics because these are not shown to the data team. We also show people the first and last name of the depositor and each potential Census match, along with their age difference, which are the typical linkage information.

We attempt to hand-link each depositor record, but also train a machine-learning model on the hand-links (see Appendix for details).[15] For links made when the machine-learning model assigns a low probability (below 0.10), we ask a person to look again at this case to confirm they think the link should be made. For links not made when the machine-learning model assigns a high probability (above 0.50), we also ask someone to look to see if this link should be made. We then re-estimate the machine-learning model on the validated data, generating a match probability that we use for robustness analysis by varying the sample based on estimated link probability thresholds.

We link 14,138 depositor records to the Census uniquely (1:1), which reflects a 32% link rate that is fairly high relative to Census-to-Census links in this time period and particularly for a sample that is predominately recently-enslaved. We attribute this higher linkage rate to our making a more contemporaneous link between Freedman Bank records and the 1870 Census, rather than linking between Censuses, and the corresponding ability to take advantage of household member names that provide high confidence in links. For example, the depositor record for Lafayette Spencer (Appendix Figure 1) corresponds closely to his Census record (Appendix Figure 2) with spouse Nancy and child Grant (that is 3 months old in the

[14]The "ambiguous" category includes the complexion categories from "Very Light Brown" to "Fair" in Table 1, along with the ambiguously named "Nearly Black."

[15]We train two random forest models, separately for child and adult links, that include 40 variables describing the relationship between the depositor and each potential census match. Some example variables reflect: string distance between first and last names, nicknames, and family member names; depositor complexion and Census race; age difference; the rarity of each first name; and the extent of missing matching fields in each record.

January Bank record and 10/12 months old in the August Census). The occupations also align (farming vs. farmer), along with the location (Hyde County, North Carolina), though we do not use occupation or location in the linkage process. Lafayette Spencer opened his account in 1870 with the New Bern branch (Craven County, North Carolina).

We restrict our analysis to these 14,138 unique links. Our main sample includes 10,286 census adults (7,268 men and 3,018 women) and 3,852 children (2,918 boys and 934 girls). The main analysis uses the pooled sample, and we see similar estimates across adults and children and by gender.

Figure 1 shows the locations of Bank branches with surviving depositor records, where the size of the circle is proportional to the number of depositors we link to the Census. Freedman branches cover the major Southern cities at the time. 54% of depositors are in their branch county and a further 8% are in a neighboring county in the 1870 Census. Our estimates are robust to restricting the analysis to depositors in the branch county or neighboring county, as more-distant people may be linked incorrectly, though people may also have opened an account further away (as in the case of Lafayette Spencer) or moved between the Census date and opening their account. Indeed, the share of depositors living in their branch county or neighboring counties falls from 71% for accounts opened in 1870 to 52% for accounts opened in 1874. As a comparison, Osthaus (1976) estimates that two-thirds of Freedman Bank depositors were in the branch city when they opened their account with select branches.

I.C 1870 Census Data

The 1870 Census is particularly useful in providing information on household wealth. Census enumerators could record a value of real estate owned and a value of personal estate owned for each person, but generally only for the household head (listed first), and we define a summed measure of household wealth. Census enumerators were instructed to not record personal wealth below \$100, and rarely record real estate wealth below \$100.[16]

We define an indicator variable for whether the depositor's household has "measured wealth," defined as $\geq$ \$100 total household wealth, which is the case for 27% of our sample households. We also analyze the continuous variation in wealth, adding \$45 for households with unmeasured wealth and taking the natural log. Our baseline adjustment assumes \$45 because some Census enumerators across the country did not follow the instructions and recorded positive values of personal wealth below \$100 (for 2% of households with less than \$100 wealth), with an average value of \$45. This value varies moderately by Census race (\$39, \$40, and \$48 for Black, Mulatto, and White household heads), and the estimated

[16]Census instructions were to record the approximate market value of real estate owned by the household, not including any mortgage, but in this era people were less likely to have mortgages and had lower debt shares on real estate than in the modern US.

impact of log wealth is very similar when assuming these race-specific values. Unmeasured household-specific wealth below $100 would generally attenuate the estimated effects of log wealth, though the estimates for log wealth are similar when restricting the sample to those with $100 or greater wealth.

The Census also reports whether people can read or write, for those aged 10 and older. People can generally do both or neither in the data, so we define people as "literate" when they can read and write. We focus on whether the depositor is literate, but also consider the literacy of the household head.

Table 2 reports summary statistics for our sample. 70% of the depositors are racialized as Black in the Census and 14% are racialized as Mulatto. Average household wealth is $770, including those with zero wealth, which reflects a highly skewed wealth distribution. Among those ages 10 and older, roughly one-third are literate. Freedman Bank depositors are moderately more likely to have measured household wealth and be literate than others of the same Census race in their county, controlling for gender and being an adult (Table 2, columns 2 – 4).

II Empirical Specifications with Race as the Outcome

Our main empirical specifications examine Census race as the outcome. We define three outcomes: whether someone is racialized as White or Mulatto (instead of Black); whether someone is racialized as White (instead of Mulatto or Black); and whether someone is racialized as Mulatto (instead of Black, excluding White people).

We regress these outcomes on depositor characteristic W_i, depositor skin tone, and branch fixed effects:

$$Race_i = \beta W_i + \delta_S SkinToneIndex_i + \gamma_b + \epsilon_i, \tag{1}$$

$$Race_i = \beta W_i + \gamma_{SkinTone} + \gamma_b + \epsilon_i. \tag{2}$$

Equation 1 uses the continuous skin tone index as a control, standardized to have a mean of zero and a standard deviation of one. Equation 2 is our main specification, however, which uses fixed effects for each of the 27 detailed skin tone categories. We include branch fixed effects, as racial composition varied across branches and different branches often used different terms for skin tones, but our estimates are not sensitive to including branch fixed effects. Similarly, while depositors opened accounts between 1865 and 1874, the estimates are not sensitive to including fixed effects for the interaction of branch and account opening year. Appendix Table 1 provides a tabulation of detailed skin tone category by branch.

II.A Main Estimates

Table 3 reports that people are 11 percentage points more likely to be racialized as White or Mulatto if they have measured wealth in the Census (Panel A, Column 1). Having a one standard deviation darker skin tone index value substantially reduces the chance that someone is racialized as White or Mulatto, by 29 percentage points, but conditional on this skin tone index there is also a substantial influence of wealth.[17] Column 2 is our preferred specification, replacing the skin tone index with detailed skin tone category fixed effects, which shows a very similar influence of wealth on racialization. Columns 3 and 4 report a 7 percentage point increase in whether people are racialized as White, compared to Mulatto or Black. Columns 5 and 6 show a 9 percentage point increase in being racialized as Mulatto compared to Black, excluding White people from the sample. Panels B and C report impacts of log wealth and literacy, in which these different proxies for socioeconomic status are positively associated with being racialized as White or Mulatto.

Table 4 reports estimates separately by broad skin tone group. As the skin tone of the depositor becomes lighter, the effects of wealth and literacy on racialization grow stronger: from 6% for those with "black" skin color to 22% for people with "light" skin tones (between "very light brown" and "fair") in Panel A of Table 4. The effect of wealth then decreases to 1% for people with "white" skin tone. Intuitively, people with light skin tones are statistical "compliers" to the influence of socioeconomic status. By contrast, when people have dark skin tones (or "white" complexion), their racialization is less sensitive to their wealth and literacy. There continues to be some smaller effect of wealth and literacy on racialization for those with darker skin tones, including "black," which can reflect more coarse or racialized skin tone measures.

Table 5 explores potential geographic heterogeneity in the effects of wealth and literacy. Column 1 reports the baseline average effect, and we then consider how these effects vary in counties with: one standard deviation lower wealth among Black people in the Census (Column 2); one standard deviation lower literacy among Black people in the Census (Column 3); and one standard deviation greater reports of racial abuses against Black people (Column 4).[18] The effects of wealth and literacy on racialization do not vary substantially across counties. There is some indication, though, that wealth matters more for how people are racialized in counties with lower Black wealth; that is, in counties where Black people

[17] A one standard deviation increase in the skin tone index corresponds to an increase of 0.3 (from Table 1), or roughly from "light" complexion to "brown" complexion.

[18] For these data on racial abuses, we are grateful to Eric Chyn, Kareem Haggag, and Bryan Stuart for sharing their data (from Chyn, Haggag and Stuart, 2024). These data detail the location of crimes reported by the Freedmen's Bureau from 1865 to 1868, historically referred to as "outrages," mostly perpetrated against African Americans. When calculating average wealth among families with Black household heads in the Census, we omit our sample families.

generally have less wealth, when someone has more wealth they are less likely to be racialized as Black. Similarly, in counties where literacy is lower among Black people, when someone is literate they are less likely to be racialized as Black. The effects of wealth are less pronounced along the Mulatto vs. Black margin, but in counties where wealth is lower among Mulatto people then people are less likely to be racialized as Mulatto when they have more wealth. In counties with more racial violence during the 1865–1868 period, there is also a somewhat greater effect of wealth on racialization, which is consistent with greater racial separation along socioeconomic lines setting the stage for racial segregation during Jim Crow in places with greater racialized violence during the Reconstruction Era.

Instead of using Census measures of wealth and literacy, we also explore the effects on Census racialization based on depositor characteristics that are recorded in the Bank data. We define names as "Distinctively Black," based on a greater prevalence of particular names among African Americans that had emerged by this era (Cook, Logan and Parman, 2014, 2016, 2022). The depositor record often lists an occupation, which is an imperfect proxy for SES due to an unknown dispersion of incomes among "farmer," "farming," "farm laborer," and "laborer," but we define categories of "Professional" and "Service" occupations (as opposed to farmers and laborers).[19] The Bank record also had a signature line, where it appears that the depositor sometimes signed their own name and sometimes the bank cashier wrote the name with an X along with writing "his mark" or "her mark." We use whether the depositor signed their own name as an imperfect proxy for literacy.

Table 6 reports impacts of these Bank record characteristics on how people were racialized in the 1870 Census. People are less likely to be racialized as White or Mulatto when they have a more distinctively Black name in their depositor record (Panel A), which suggests additional channels through which race is socially constructed that need not be directly associated with socioeconomic status. Relatedly, for families with multiple depositors, 79% of depositors are in the same coarse skin tone group but 93% of these depositors are assigned the same race in the Census – consistent with Census enumerators characterizing a person's race based also on their family.[20] If depositors are listed as having a professional or service job, they are more likely to be racialized as White or Mulatto (Panel B), and also if they

[19]Examples of Professional occupations include: Barber, Nurse, Teacher. Examples of Service occupations include: Cook, Drayman, Waiter. The omitted comparison category includes farmers and/or laborers, which are difficult to distinguish in these records at this time. We exclude those with missing occupations or students (i.e., children in school). The occupational composition of depositors is: 1,126 "professional," 3,530 "service," 5,790 "farmer and/or laborer," and 3,692 missing or uncategorized.

[20]There are 680 families in which at least one depositor is not in the "white" skin tone category. Among these families, 93% have the same Census race assigned to each depositor (Black, Mulatto, White) whereas 79% have the same coarse skin tone group (Darker, Light, White). The "Light" category corresponds to the detailed skin tone categories between "Fair" and "Very Light Brown" in Table 1.

wrote their name on the signature line (Panel C). These magnitudes are smaller, particularly for being racialized as White (rather than Mulatto or Black), but generally consistent with Census enumerators racializing people differently based on characteristics from the depositor records after controlling for detailed skin tone complexion fixed effects.

II.B Unobserved Socioeconomic and Physical Characteristics

We emphasize that how people were racialized was strongly associated with their wealth, conditional on their physical skin tone, which indicates the social construction of race along economic lines apart from physical differences between people. Wealth is a proxy for socioeconomic status, more generally, and the interpretation remains the same when racialization is also influenced by other socioeconomic characteristics that are correlated with wealth. These other socioeconomic characteristics include observed factors, such as literacy, and unobserved factors like how people dress or speak. We are not saying that an exogenous increase in wealth would itself necessarily cause someone to be racialized differently; rather, observed wealth captures a bundle of socioeconomic characteristics that cause someone to be racialized differently.

The interpretation would be different, however, if race was entirely determined by other unobserved physical characteristics that are correlated with wealth. These other physical characteristics might reflect variation in skin tone that is not captured by measured skin tone, or other physical features that people may use when racializing themselves and others (such as hair texture or facial features). To the extent that unobserved physical features predict race, and are correlated with wealth, one could try to maintain the null hypothesis that race was entirely determined by physical features and not constructed along socioeconomic lines.[21]

We have an opportunity to bound the influence of unobserved physical characteristics, using the influence of detailed skin tone data that we do observe. This bounding exercise follows Cinelli and Hazlett (2020), related to Altonji, Elder and Taber (2005) and Oster (2019). If we only observed three coarse groupings of skin tone (White, Light, Darker), similar to what is more often observed when skin tone data are available, then we could still look to estimate the effect of wealth on racialization conditional on those coarse skin tone group fixed effects.[22] We would be concerned about unobserved physical characteristics, however, because of potential remaining variation in skin tone (very light vs. light vs. light brown and brown vs. dark brown vs. dark). When we augment the analysis with more detailed skin

[21] Physical features can also be endogenous to wealth and socioeconomic status, but here we consider the potential concern in which physical features are exogenous and predictive of race (while being correlated with wealth).

[22] For this specification, we group the Table 1 categories into: White, Light (from "Fair" to "Very Light Brown"), and Darker (from "Dark Mulatto" to "Black").

tone data, we can quantify how including the skin tone index affects the estimates (due to its joint predictive power for race and its correlation with wealth beyond the three coarse group fixed effects). We can then consider how our estimated effects would change if we discovered yet another measure of physical characteristics, which was equally jointly predictive of race and wealth as the detailed skin tone data.

Figure 3 shows the limited scope for unobserved physical characteristics to explain our main estimated impacts of wealth on racialization. The closest red diamond indicates that the estimated effect of wealth would drop from 0.109 to 0.104 with the addition of another variable that was as influential as adding the skin tone index.[23] With the addition of another summary measure of physical characteristics five times as influential as the skin tone index, the estimated effect of wealth would drop to 0.080 (the second red diamond). To entirely explain the effect of wealth, reaching zero at the dotted blue line, unobserved physical characteristics would need to be 19 times as correlated with race and wealth as the detailed data on skin tone – which itself was the primary identifying information collected by the Bank.

In Figure 3, the curved contour lines represent adjusted effects of wealth on race due to unobserved characteristics having the indicated joint explanatory power for race (on the y-axis) and wealth (on the x-axis). If unobserved physical characteristics explained all residual variation in race, moving up the y-axis to one, then these characteristics would still need to be 3.6 times more correlated with wealth than the skin tone index.

Figure 4 shows similar robustness exercises for the effects of log wealth and literacy on whether people are racialized as White or Mulatto rather than Black. Figure 5 shows robustness for the effects of measured wealth on being White rather than Black or Mulatto, which is more sensitive but still requires unobserved characteristics to be substantially more explanatory than our detailed skin tone data, and shows a more robust effect of measured wealth on whether people are racialized as Mulatto compared to Black.

Overall, these estimates indicate a substantial and robust relationship between proxies for socioeconomic status and racialization that would be difficult to account for quantitatively with unobserved physical characteristics of depositors. These estimates show that in 1870, shortly after emancipation, racial groups were being separated along socioeconomic lines in setting the stage for racial segregation and continued racial stratification.

II.C Robustness to Linkage Error and Alternative Specifications

Record linkage error is a main source of concern that we look to address. Because White and Mulatto people tend to have more wealth or higher literacy than Black people in the Census, inaccurate links to White or Mulatto people could cause us to observe people that both have

[23]The baseline effect is normalized to the origin, which corresponds to the specification from Table 3 (Panel A, Column 2) with the additional coarse skin tone group fixed effects.

higher wealth or literacy and are being racialized as White or Mulatto. Linkage errors could then generate spurious "effects," but would not generate the differentially greater effect of wealth seen in Table 4 among those with light skin tone (and our data team did not observe wealth or literacy when linking records). Linkage error would also attenuate estimates in our specifications using depositors' Bank record characteristics.

Appendix Table 2 shows that our estimates are robust to restricting the record links included in the analysis. Column 2 excludes links that the machine-learning model assigns a probability below 0.10, and Column 3 excludes links that the machine-learning model assigns a probability below 0.50. Our preferred conservative approach more directly leverages the other household member names, however, restricting the sample to depositors with at least one other family member whose name matches (Column 4) or at least two other family members whose names match (Column 5). Column 6 finds similar estimates when restricting our hand-linked sample to depositors whose name matches exactly to the Census name, and their age difference is within one year (due to variation in birthdays relative to the Census date and Bank account opening date). Appendix Table 3 shows similar effects of depositors' Bank record characteristics when making the same sample restrictions. We continue to find similar effects on racialization even as restricted samples might become less representative of the broader population (e.g., able to link on spouse and child names).

Appendix Table 4 reports estimates when restricting the sample to depositors living in the same branch county or neighboring counties as when they opened their Bank account (Column 2). We do not link on location, and these links may be more reliable, but record links across locations are often verified with names of multiple other household members (as in the case of Lafayette Spencer). The estimates in this restricted sample are smaller in magnitude, but migration is endogenous and potentially associated with depositors' other economic characteristics and how they are racialized in the Census. The estimates are similar to our baseline estimates when excluding the New York City branch (Column 3) and excluding all foreign-born people in the Census (Column 4).

Appendix Table 5, Column 2, reports larger estimates when omitting the more "racial" categories of skin tone (black, nearly black, colored, mulatto, white, nearly white). These "racial" skin tone labels may be particularly subject to measurement error if some Freedman Bank cashiers, like Census enumerators, were susceptible to write down "black" or darker skin tones when people looked poorer. This behavior would attenuate our baseline estimates: depositors would have lighter racial skin tone categories recorded when they are richer, even though their skin tone is darker, so that within measured skin tone categories wealthier people would be *less* likely to be racialized as White or Mulatto by the Census (because they actually have darker skin tones).

Column 3 of Appendix Table 5 limits the sample to Household Heads only. It is not indicated who was present with the Census enumerator, but Appendix Table 6 shows that the literacy of the Household Head is similar or more influential than the literacy of the depositor themselves.

Appendix Table 5 also shows similar effects across gender and age, restricting the sample to women (Column 4) and children (Column 5). Column 6 restricts the sample to households with measured wealth ($100 or greater), which shows similar effects of log wealth and somewhat larger variation with literacy. Appendix Tables 7 and 8 separate the influence of real estate wealth and personal property wealth, in which both matter individually and jointly, but having real estate wealth is somewhat more influential.

Appendix Table 9 reports similar estimates when varying the included control variables. Column 2 omits the branch fixed effects. Columns 3 – 6 include additional controls for: fixed effects for branch by year of account opening, being emancipated prior to 1860, gender along with age and age-squared, and county fixed effects.[24]

III Empirical Specifications with Race as an Explanatory Variable

One implication of race being influenced by socioeconomic status, which we develop in this section, is that it complicates the interpretation of empirical specifications that "control" for endogenous race. To illustrate this, we start by regressing depositor characteristic W_i on the skin tone index:

$$W_i = \beta_S SkinToneIndex_i + \gamma_b + \epsilon_i, \tag{3}$$

where β_S can be interpreted as the average difference for depositors with one standard deviation darker skin tone. We continue to control for Freedman branch fixed effects, so this coefficient is the average difference among depositors associated with the same Bank branch.

Table 7, Column 1, reports that depositors with one standard deviation darker skin tone are 12 percentage points less likely to have measured wealth, have 41% lower household wealth, and are 17 percentage points less likely to be literate.

A common empirical specification would then "control for race," examining the influence of skin tone within race, by including indicator variables for being Black and Mulatto in the Census:

$$Y_i = \tilde{\beta}_S SkinToneIndex_i + \tilde{\beta}_B Black_i + \tilde{\beta}_M Mulatto_i + \gamma_b + \epsilon_i. \tag{4}$$

[24]As an imperfect proxy for whether people were free in 1860, we use a machine-learning model to link Black and Mulatto people from 1870 back to 1860 when free people were enumerated by name in the Census (and enslaved people were counted without names).

Table 7, Column 2, reports substantially smaller estimated differences for depositors with one standard deviation darker skin tone: 1 percentage point less likely to have measured wealth, 10% lower household wealth, and 6 percentage points less likely to be literate. Columns 3 – 5 report smaller differences by skin tone within each racial group, which is also a typical comparison made.

These estimates might be interpreted as a smaller true relationship between wealth and skin tone, but we simulate and show empirically that these estimates in columns 2 – 5 are biased by the endogenous assignment of race and understate the relationship with skin tone. For the relationship between other outcomes and skin tone, bias from the endogenous assignment of race need not be so large.[25] But we show below that for a fixed known relationship between wealth and skin tone, the estimated effect can be substantially attenuated when racial assignment is influenced by wealth.

Intuitively, despite an overall relationship between darker skin tones and lower wealth, the observed relationship will be attenuated both among people considered Black and among people considered White. People who are "Black" with light skin will be relatively poor because otherwise they would not be considered Black, so light skin becomes less associated with higher wealth. Relatedly, people who are "White" with dark skin will be relatively rich because otherwise they would not be considered White, so dark skin becomes less associated with lower wealth. For conciseness, we focus on "Black" and "White" and omit a third "Mulatto" category that would introduce additional notation below with the same implications.

Selection bias from the endogenous assignment of race can apply more broadly than in the case of estimating relationships with skin tone (as in Table 7). Regression of any outcome on a predictor, in which the predictor has a causal effect on the assignment of individuals to race, will be subject to this type of bias when controlling for race. For example, this includes specifications such as a regression of wealth on education controlling for race. In this specification, the effect of education is attenuated with the inclusion of race controls when, as we formalize below, individuals with lower earnings are more likely to be assigned the "Black" category.

[25]Appendix Table 10 shows that a darker skin tone is negatively associated with having a professional occupation in the Bank record and signing the depositor book (and positively associated with having a more distinctively Black name). This relationship is less attenuated, after controlling for race, which is consistent with Census race being less endogenous to these measures. It is also consistent with less variation in these outcomes by racial category and more variation with skin tone.

III.A Modeling Selection Bias from Endogenous Race

Let w_i be the wealth of person i, assumed to be decreasing in darker skin tone s_i:

$$w_i = \alpha_0 + \alpha_1 s_i + e_i. \tag{5}$$

If s_i is measured accurately, then an OLS regression of w_i on s_i yields the true α_1. Later, we consider measurement error in s_i.

Suppose we estimate the same regression among respondents classified as Black:

$$w_i = \tilde{\alpha}_0 + \tilde{\alpha}_1 s_i + e_i, \text{ if } B_i = 1. \tag{6}$$

If racial classification is only a function of skin tone, then $\tilde{\alpha_1} = \alpha_1$. However, if the probability of classification as Black increases with other characteristics u_i, such as poverty or illiteracy with $\text{Cov}(u_i, e_i) < 0$, then $\tilde{\alpha}_1$ is biased upwards ($\tilde{\alpha}_1 > \alpha_1$ and $\alpha_1 < 0$).

When people are classified as Black based on skin tone and these other factors u_i:

$$B_i = \mathbb{I}\left[\beta_0 + \beta_1 s_i + u_i > 0\right], \text{with } \beta_1 > 0. \tag{7}$$

Wealth within the group classified as Black is:

$$\mathbb{E}\left[w_i | Black\right] = \tilde{\alpha}_0 + \tilde{\alpha}_1 s_i + \mathbb{E}\left[e_i | u_i > -\left(\beta_0 + \beta_1 s_i\right)\right], \tag{8}$$

and the error term is no longer orthogonal to skin tone. To illustrate this bias in a tractable case, we assume the errors are distributed normally ($e_i \sim N\left(0, \sigma_e^2\right)$, $u_i \sim N\left(\mu_u, \sigma_u^2\right)$) and negatively correlated ($\text{Corr}(u_i, e_i) = \rho < 0$). Then, expected wealth conditional on being classified as Black is:

$$\mathbb{E}\left[w_i | Black\right] = \tilde{\alpha}_0 + \tilde{\alpha}_1 s_i + \rho\sigma_u I\left(\beta_0 + \beta_1 s_i\right), \tag{9}$$

where $I(\cdot) = \phi(\cdot)/\Phi(\cdot)$ is the inverse Mills ratio (Heckman, 1979). This final term is unobserved, creating selection bias in α:

$$\tilde{\alpha}_1 = \alpha_1 + \rho\sigma_u \frac{Cov\left(s_i, I_i\right)}{Var\left(s_i\right)}. \tag{10}$$

Compared to the true α_1, which is less than zero, the estimated $\tilde{\alpha}_1$ is biased upwards toward zero or even positive. Because signals of poverty or illiteracy are negatively correlated with wealth ($\rho < 0$) and selected racial assignment makes the inverse Mills ratio lower with darker

skin tones ($\partial I_i/\partial s_i < 0$ hence $Cov(s_i, I_i) < 0$), the second term is positive and $\tilde{\alpha}_1$ is biased upwards ($\tilde{\alpha}_1 > \alpha_1$).

III.B Simulated and Empirical Correlations between Wealth and Skin Tone, Unconditional and Conditional on Census Race

Figure 6, Panel A, shows simulated correlations between wealth and skin tone, based on the above selection model with assumed illustrative parameters.[26] The hollow circles represent people racialized as White, whereas the black circles represent people racialized as Black.

Unconditional on Census race, the solid red line shows the assumed decreasing relationship between wealth and darker skin tone. Conditional on race, however, the two dashed lines show substantially attenuated relationships. Among those with darker skin, only richer people are White, which flattens the top dashed line. Among those with lighter skin, only poorer people are Black, which flattens the bottom dashed line.

Empirically, in our data, we see a similar pattern in Panel B of Figure 6. We plot log household wealth on the y-axis against the depositor's skin tone index on the x-axis.[27] The solid red line shows a clear unconditional decline in wealth associated with darker skin tones. Conditional on race, however, the two dashed red lines show substantially flatter relationships. This is because among those with light skin tones, people racialized as Black are relatively poorer on average. Similarly, among those with darker skin tones, people racialized as White are relatively richer on average. This selected racialization creates upward bias within both racial groups.

III.C Measurement Error in Skin Tone

We can also consider implications if the Freedman Bank cashiers recorded skin tone based partly upon signals of socioeconomic status, v_i:

$$\hat{s}_i = s_i + v_i. \tag{11}$$

An OLS regression of wealth on skin tone becomes:

$$w_i = \hat{\alpha}_o + \hat{\alpha}_1 (s_i + v_i) + \eta_i \tag{12}$$

and the estimated coefficient $\hat{\alpha}_1 = \beta_{e\hat{s}} + (1 - \beta_{v\hat{s}})\,\alpha_1$.

The first term, $\beta_{e\hat{s}}$, creates a downward bias in the relationship between wealth and skin

[26] We set $w_i = -0.5 \times s_i + e$, with s randomly drawn from a standard normal distribution, and let $B_i = 1$ if $0.75 \times s_i + u > -0.5$. We draw 500 error terms, e and u, from a standard normal distribution with a -0.75 correlation. We then fit population regression lines to the full sample and within each racial group.

[27] We restrict the sample to people with measured household wealth to aid in visualizing this relationship, as the bottom-coded households are grouped together.

tone (more negative) when the Bank's measurement of skin tone is affected by signals of wealth. The coefficient $\beta_{e\hat{s}}$ is from a regression of e on $\hat{s}$, which is negative ($\text{Corr}(v_i, e_i) < 0$). The same type of bias would also arise when someone's measured skin tone actually becomes darker due to lower wealth, such as working as an outdoors laborer in a lower-income job, which creates a more negative relationship between measured skin tone and wealth.

The second term represents attenuation bias, as with classical measurement error, where $\beta_{v\hat{s}}$ is the coefficient from a regression of v on $\hat{s}$ (positive). This attenuation bias could be exacerbated when controlling for Census race, also reducing the coefficients in column 2 relative to column 1.

To assess the attenuation from classical measurement error, we can use the two measures of skin tone from when some depositors opened two accounts at different times. Table 8 restricts our analysis to this subsample and finds 2SLS estimates that are similar to the OLS estimates, instrumenting one measurement of skin tone with the other. The first-stages are very strong, unconditional or conditional on Census race, with F-statistics of 1,626 and 437 due to high correlation in the multiple measures of skin tone. These estimates indicate little attenuation bias from classical measurement error, but substantial selection bias from the endogenous assignment of race based on wealth.

IV Conclusion

Emancipation put the United States' social structure into a period of potential transformation. The Reconstruction Era saw efforts to integrate formerly-enslaved people into broader social, economic, and political institutions. When Reconstruction collapsed, however, and there was a backlash movement to reimpose antebellum social structures, these efforts had to confront the continuous variation in physical skin tones across people.

Without the de jure distinction between those who were to be considered "free" versus "enslaved," the definition of race became an even more important tool for social control. Through the Jim Crow era of segregation, racial categories were a central pillar of the social and economic system, both reflecting and influencing the beliefs and behaviors of individuals in that system. This required a social construction of race, particularly in that critical juncture after emancipation, which set the stage for continued racial stratification.

In this paper, we show that people with the same physical skin tone were racialized differently based on their wealth and other proxies for socioeconomic status. We use uniquely-detailed data on skin tone from depositor records of the Freedman's Savings Bank (1865-1874) to show differential racialization by socioeconomic status in the 1870 Census. The influence of socioeconomic status on racialization is similar across age, gender, and geographical area, suggesting a pervasive social construction of race even after the upheaval of

emancipation.

Standard statistical practice still often involves "controlling for race" alongside other variables, despite the growing recognition that "race is a social construct." When those racial categories themselves are outcomes of the analysis, analyzing outcome differences across or within racial groups can be challenging to interpret. For example, observed differences by skin color within race are attenuated when race is endogenous.

Our findings underscore the fluidity of socially constructed racial classifications, which challenge essentialist views of racial identity that continue to influence both the academic literature and social structure. Our estimates also not only reinforce the idea that race is a social construct, but show empirically the historical construction of race in this critical moment in American history after emancipation as a new era of social coercion emerged.

References

Abramitzky, Ran, Jacob Conway, Roy Mill, and Luke Stein. 2023. "The Gendered Impacts of Perceived Skin Tone: Evidence from African-American Siblings in 1870–1940." NBER Working Paper 31016.

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, James Feigenbaum, and Santiago Pérez. 2021. "Automated Linking of Historical Data." *Journal of Economic Literature*, 59(3): 865–918.

Alba, Richard, and Tariqul Islam. 2009. "The Case of the Disappearing Mexican Americans: An Ethnic-Identity Mystery." *Population Research and Policy Review*, 28(2): 109–121.

Alba, Richard D., and Mitchell B. Chamlin. 1983. "A Preliminary Examination of Ethnic Identification among Whites." *American Sociological Review*, 48(2): 240–247.

Althoff, Lukas, and Hugo Reichardt. 2024. "Jim Crow and Black Economic Progress after Slavery." *The Quarterly Journal of Economics*, 139(4): 2279–2330.

Altonji, Joseph G., Todd E. Elder, and Christopher R Taber. 2005. "Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113(1): 151–184.

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

Baron, E. Jason, Joseph J. Doyle Jr., Natalia Emanuel, Peter Hull, and Joseph P. Ryan. 2024. "Unwarranted Disparity in High-Stakes Decisions: Race Measurement and Policy Responses." *NBER Working Paper 33104.*

Blake, Jamilia, Verna Keith, Wen Luo, Huong Le, and Phia Salter. 2017. "The Role of Colorism in Explaining African American Females' Suspension Risk." *School Psychology Quarterly*, 32(1): 118–130.

Boas, Franz. 1912. "Changes in the Bodily Form of Descendants of Immigrants." *American Anthropologist*, 14(3): 530–562.

Bodenhorn, Howard. 2015. *The Color Factor: The Economics of African-American Well-Being in the Nineteenth-Century South.* Oxford University Press.

Bonilla-Silva, Eduardo. 1997. "Rethinking Racism: Toward a Structural Interpretation." *American Sociological Review*, 62(3): 465–480.

Bowman, Phillip J., Ray Muhammad, and Mosi Ifatunji. 2004. "Skin Tone, Class, and Racial Attitudes among African Americans." In *Skin Deep: How Race and Complexion Matter in the "Color-blind" Era.* 128–58. University of Illinois Press Chicago.

Burch, Traci. 2015. "Skin Color and the Criminal Justice System: Beyond Black-White

Disparities in Sentencing." *Journal of Empirical Legal Studies*, 12(3): 395–420.

Charles, Kerwin K., and Jonathan Guryan. 2011. "Studying Discrimination: Fundamental Challenges and Recent Progress." *Annual Review of Economics*, 3(1): 479–511.

Chelwa, Grieve, Darrick Hamilton, and James Stewart. 2022. "Stratification Economics: Core Constructs and Policy Implications." *Journal of Economic Literature*, 60(2): 377–399.

Chyn, Eric, Kareem Haggag, and Bryan A. Stuart. 2022. "The Effects of Racial Segregation on Intergenerational Mobility: Evidence from Historical Railroad Placement." *NBER Working Paper 30563*.

Chyn, Eric, Kareem Haggag, and Bryan A. Stuart. 2024. "Inequality and Racial Backlash: Evidence from the Reconstruction Era and the Freedmen's Bureau." NBER Working Paper 32314.

Cinelli, Carlos, and Chad Hazlett. 2020. "Making Sense of Sensitivity: Extending Omitted Variable Bias." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(1): 39–67.

Cook, Lisa D., Trevon D. Logan, and John M. Parman. 2014. "Distinctively Black Names in the American Past." *Explorations in Economic History*, 53: 64–82.

Cook, Lisa D., Trevon D. Logan, and John M. Parman. 2016. "The Mortality Consequences of Distinctively Black Names." *Explorations in Economic History*, 59: 114–125.

Cook, Lisa D., Trevon D. Logan, and John M. Parman. 2022. "The Antebellum Roots of Distinctively Black Names." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 55(1): 1–11.

Célérier, Claire, and Purnoor Tak. 2024. "Finance, Advertising, and Race." *Review of Financial Studies*. Forthcoming.

Dahis, Ricardo, Emily Nix, and Nancy Qian. 2020. "Choosing Racial Identity in the United States, 1880-1940." NBER Working Paper 27774.

Darity Jr., William A. 2005. "Stratification Economics: The Role of Intergroup Inequality." *Journal of Economics and Finance*, 29(2): 144–153.

Darity Jr., William A., Darrick Hamilton, and James B. Stewart. 2015. "A Tour de Force in Understanding Intergroup Inequality: An Introduction to Stratification Economics." *The Review of Black Political Economy*, 42(1-2): 1–6.

Darity Jr., William A., Patrick L. Mason, and James B. Stewart. 2006. "The Economics of Identity: The Origin and Persistence of Racial Identity Norms." *Journal of Economic Behavior & Organization*, 60(3): 283–305.

Davenport, Lauren. 2020. "The Fluidity of Racial Classifications." *Annual Review of Political Science*, 23: 221–240.

Davis, F. James. 2010. *Who Is Black?: One Nation's Definition.* Penn State Press.

Derenoncourt, Ellora, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick. 2024. "Wealth of Two Nations: The U.S. Racial Wealth Gap, 1860–2020." *The Quarterly Journal of Economics*, 139(2): 693–750.

Diette, Timothy M., Arthur H. Goldsmith, Darrick Hamilton, and William Jr. Darity. 2015. "Skin Shade Stratification and the Psychological Cost of Unemployment: Is there a Gradient for Black Females?" *The Review of Black Political Economy*, 42(1): 155–177.

Dixon, Angela R., and Edward E. Telles. 2017. "Skin Color and Colorism: Global Research, Concepts, and Measurement." *Annual Review of Sociology*, 43(1): 405–424.

Douglass, Frederick. 1874. "Letter to Gerrit Smith."

DuBois, W.E.B. 1903. *The Souls of Black Folk.* A.C. McClurg & Co.

Duncan, Brian, and Stephen J. Trejo. 2011. "Intermarriage and the Intergenerational Transmission of Ethnic Identity and Human Capital for Mexican American." *Journal of Labor Economics*, 29(2): 195–227.

Duncan, Brian, and Stephen J. Trejo. 2017. "The Complexity of Immigrant Generations: Implications for Assessing the Socioeconomic Integration of Hispanics and Asians." *ILR Review*, 70(5): 1146–1175.

Duncan, Brian, and Stephen J. Trejo. 2023. "Which Mexicans are White? Enumerator-assigned Race in the 1930 Census and the Socioeconomic Integration of Mexican Americans." *NBER Working Paper 31623.*

Duncan, Brian, Jeffrey Grogger, Ana Sofia Leon, and Stephen J. Trejo. 2020. "New Evidence of Generational Progress for Mexican Americans." *Labour Economics*, 62: 101771.

Engerman, Stanley L. 2003. "Slavery Without Racism, Racism Without Slavery: Mainland North America and Elsewhere." Yale University. New Haven, Connecticut. Conference held November 7-8, 2003.

Eschbach, Karl, Khalil Supple, and C. Matthew Snipp. 1998. "Changes in Racial Identification and the Educational Attainment of American Indians, 1970–1990." *Demography*, 35(1): 35–43.

Finlay, Keith, Elizabeth Luh, and Michael G. Mueller-Smith. 2024. "Race and Ethnicity (Mis)measurement in the U.S. Criminal Justice System." *NBER Working Paper 32657.*

Francis, Dania V., Bradley L. Hardy, and Damon Jones. 2022. "Black Economists on Race and Policy: Contributions to Education, Poverty and Mobility, and Public Finance." *Journal of Economic Literature*, 60(2): 454–493.

Freeman, Jonathan B., Andrew M. Penner, Aliya Saperstein, Matthias Scheutz,

and Nalini Ambady. 2011. "Looking the Part: Social Status Cues Shape Race Perception." *PloS ONE*, 6(9): e25107.

Goldsmith, Arthur H., Darrick Hamilton, and William A. Darity Jr. 2006. "Shades of Discrimination: Skin Tone and Wages." *American Economic Review*, 96(2): 242–245.

Goldsmith, Arthur H., Darrick Hamilton, and William A. Darity Jr. 2007. "From Dark to Light: Skin Color and Wages Among African-Americans." *The Journal of Human Resources*, 42(4): 701–738.

Gyimah-Brempong, Kwabena, and Gregory N. Price. 2006. "Crime and Punishment: and Skin Hue Too?" *American Economic Review*, 96(2): 246–250.

Hahn, Robert A., Joseph Mulinare, and Steven M. Teutsch. 1992. "Inconsistencies in Coding of Race and Ethnicity Between Birth and Death in US Infants: A New Look at Infant Mortality, 1983 through 1985." *JAMA*, 267(2): 259–263.

Haney López, Ian F. 1994. "The Social Construction of Race: Some Observations on Illusion, Fabrication, and Choice." *Harvard Civil Rights-Civil Liberties Law Review*, 29: 1–62.

Harburg, Ernest, Lillian Gleibermann, Peter Roeper, M. Anthony Schork, and William J. Schull. 1978. "Skin Color, Ethnicity, and Blood Pressure I: Detroit Blacks." *American Journal of Public Health*, 68(12): 1177–1183.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153–161.

Hersch, Joni. 2006. "Skin-Tone Effects among African Americans: Perceptions and Reality." *American Economic Review*, 96(2): 251–255.

Hersch, Joni. 2018. "Colorism Against Legal Immigrants to the United States." *American Behavioral Scientist*, 62(14): 2117–2132.

Hirschfeld, Lawrence A. 2008. "Children's Developing Conceptions of Race." *Handbook of Race, Racism, and the Developing Child*, 37–54.

Hochschild, Jennifer L., and Brenna M. Powell. 2008. "Racial Reorganization and the United States Census 1850-1930: Mulattoes, Half-Breeds, Mixed Parentage, Hindoos, and the Mexican Race." *Studies in American Political Development*, 22(1): 59–96.

Hughes, Michael, and Bradley R. Hertel. 1990. "The Significance of Color Remains: A Study of Life Chances, Mate Selection, and Ethnic Consciousness among Black Americans." *Social Forces*, 68(4): 1105–1120.

Humes, Karen, and Howard Hogan. 2009. "Measurement of Race and Ethnicity in a Changing, Multicultural America." *Race and Social Problems*, 1: 111–131.

Hunter, Margaret. 2016. "Colorism in the Classroom: Skin Tone Stratifies African American and Latina/o Students." *Theory Into Practice*, 55(1): 54–61.

Ioannidis, John P. A., Neil R. Powe, and Clyde W. Yancy. 2021. "Recalibrating the Use of Race in Medical Research." *JAMA*, 325(7): 623–624.

Johnson, Charles. 1934. *Shadow of the Plantation.* University of Chicago Press.

Johnson, James H., Elisa Jayne Bienenstock, and Jennifer A. Stoloff. 1995. "An Empirical Test of the Cultural Capital Hypothesis." *The Review of Black Political Economy*, 23(4): 7–27.

Katz, Phyllis A., and Jennifer A. Kofkin. 1997. "Race, Gender, And Young Children." In *Developmental Psychopathology: Perspectives On Adjustment, Risk, And Disorder*. 51–74. Cambridge University Press.

Keith, Verna M., and Cedric Herring. 1991. "Skin Tone and Stratification in the Black Community." *American Journal of Sociology*, 97(3): 760–778.

Kessen, William, and Sally Weiskopf. 1976. "Color Vision and Hue Categorization in Young Human Infants." *Journal of Experimental Psychology: Human Perception and Performance*, 2(1): 115–129.

Kinzler, Katherine D., and Elizabeth S. Spelke. 2011. "Do Infants Show Social Preferences for People Differing in Race?" *Cognition*, 119(1): 1–9.

Klonoff, Elizabeth A., and Hope Landrine. 2000. "Is Skin Color a Marker for Racial Discrimination? Explaining the Skin Color–Hypertension Relationship." *Journal of Behavioral Medicine*, 23(4): 329–338.

Kosack, Edward, and Zachary Ward. 2020. "El Sueño Americano? The Generational Progress of Mexican Americans Prior to World War II." *The Journal of Economic History*, 80(4): 961–995.

Kreisman, Daniel, and Marcos A. Rangel. 2015. "On the Blurring of the Color Line: Wages and Employment for Black Males of Different Skin Tones." *The Review of Economics and Statistics*, 97(1): 1–13.

Krieger, Nancy, Stephen Sidney, and Eugenie Coakley. 1998. "Racial Discrimination and Skin Color in the CARDIA Study: Implications for Public Health Research. Coronary Artery Risk Development in Young Adults." *American Journal of Public Health*, 88(9): 1308–1313.

Landor, Antoinette M., Leslie Gordon Simons, Ronald L. Simons, Gene H. Brody, Chalandra M. Bryant, Frederick X. Gibbons, Ellen M. Granberg, and Janet N. Melby. 2013. "Exploring the Impact of Skin Tone on Family Dynamics and Race-Related Outcomes." *Journal of Family Psychology*, 27(5): 817–826.

Laudares, Humberto, and Felipe Valencia Caicedo. 2023. "Tordesillas, Slavery and the Origins of Brazilian Inequality." *CEPR Discussion Paper 18245*.

Lieberson, Stanley, and Mary C. Waters. 1993. "The Ethnic Responses of Whites:

What Causes Their Instability, Simplification, and Inconsistency?" *Social Forces*, 72(2): 421–450.

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

Louie, Patricia, and Rima Wilkes. 2018. "Representations of Race and Skin Tone in Medical Textbook Imagery." *Social Science & Medicine*, 202: 38–42.

Lovell, Peggy A., and Charles H. Wood. 1998. "Skin Color, Racial Identity, and Life Chances in Brazil." *Latin American Perspectives*, 25(3): 90–109.

Luh, Elizabeth. 2023. "Not so Black and White: Uncovering Racial Bias Through Systematically Misreported Trooper Reports." *Working Paper*.

Majumdar, Rochona. 2004. "Looking for Brides and Grooms: Ghataks, Matrimnials, and the Marriage Market in Colonial Calcutts, circa 1875–1940." *The Journal of Asian Studies*, 63(4): 911–935.

McGee, Dan. 2025. "Exploitation Through Racialization." *The Quarterly Journal of Economics*, Forthcoming.

Mill, Roy, and Luke C.D. Stein. 2016. "Race, Skin Color, and Economic Outcomes in Early Twentieth-Century America." *Working Paper*.

Monk Jr., Ellis P. 2014. "Skin Tone Stratification Among Black Americans, 2001–2003." *Social Forces*, 92(4): 1313–1337.

Monk Jr., Ellis P. 2016. "The Consequences of "Race and Color" in Brazil." *Social Problems*, 63(3): 413–430.

Montagu, Ashley. 1945. *Man's Most Dangerous Myth: The Fallacy of Race.* Columbia University Press.

Morning, Ann. 2011. *The Nature of Race: How Scientists Think and Teach about Human Difference.* University of California Press.

Noghanibehambari, Hamid, and Jason Fletcher. 2025. "Passing as White: Racial Identity and Old-Age Longevity." NBER Working Paper 33394.

Omi, Michael, and Howard Winant. 2014. *Racial Formation in the United States.* Routledge.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics*, 37(2): 187–204.

Osthaus, Carl R. 1976. *Freedmen, Philanthropy, and Fraud: A History of the Freedman's Savings Bank.* University of Illinois Press.

Penner, Andrew M., and Aliya Saperstein. 2008. "How Social Status Shapes Race." *Proceedings of the National Academy of Sciences*, 105(50): 19628–19630.

Penner, Andrew M., and Aliya Saperstein. 2013. "Engendering Racial Perceptions: An

Intersectional Analysis of How Social Status Shapes Race." *Gender & Society*, 27(3): 319–344.

Rangel, Marcos A. 2015. "Is Parental Love Colorblind? Human Capital Accumulation within Mixed Families." *The Review of Black Political Economy*, 42(1): 57–86.

Ransford, H. Edward. 1970. "Skin Color, Life Chances, and Anti-White Attitudes." *Social Problems*, 18(2): 164–179.

Reece, Robert L. 2018. "Genesis of U.S. Colorism and Skin Tone Stratification: Slavery, Freedom, and Mulatto-Black Occupational Inequality in the Late 19th Century." *The Review of Black Political Economy*, 45(1): 3–21.

Reece, Robert L. 2019. "Color Crit: Critical Race Theory and the History and Future of Colorism in the United States." *Journal of Black Studies*, 50(1): 3–25.

Reece, Robert L. 2021. "The Missing Social Movement: Colorism in Black America." *Interface: A Journal for and about Social Movements*, 13(2): 293–311.

Rogers, Joel Augustus. 1952. *Nature Knows No Color-Line.* Self-published.

Rose, Evan K. 2023. "A Constructivist Perspective on Empirical Discrimination Research." *Journal of Economic Literature*, 61(3): 906–923.

Saperstein, Aliya, and Aaron Gullickson. 2013. "A "Mulatto Escape Hatch" in the United States? Examining Evidence of Racial and Social Mobility During the Jim Crow Era." *Demography*, 50(5): 1921–1942.

Saperstein, Aliya, and Andrew M. Penner. 2012. "Racial Fluidity and Inequality in the United States." *American Journal of Sociology*, 118(3): 676–727.

Saperstein, Aliya, and Andrew M. Penner. 2016. "Still Searching for a True Race? Reply to Kramer et al. and Alba et al." *American Journal of Sociology*, 122(1): 263–285.

Schwartzman, Luisa Farah. 2007. "Does Money Whiten? Intergenerational Changes in Racial Classification in Brazil." *American Sociological Review*, 72(6): 940–963.

Seltzer, Richard, and Robert C. Smith. 1991. "Color Differences in the Afro-American Community and the Differences they Make." *Journal of Black Studies*, 21(3): 279–286.

Sen, Maya, and Omar Wasow. 2016. "Race as a 'Bundle of Sticks': Designs that Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science*, 19: 499–522.

Smedley, Audrey, and Brian D. Smedley. 2005. "Race as Biology is Fiction, Racism as a Social Problem is Real: Anthropological and Historical Perspectives on the Social Construction of Race." *American Psychologist*, 60(1): 16–26.

Telles, Edward, and Tianna Paschel. 2014. "Who Is Black, White, or Mixed Race? How Skin Color, Status, and Nation Shape Racial Classification in Latin America." *American Journal of Sociology*, 120(3): 864–907.

U.S. Census Office. 1870. "Instructions to Assistant Marshals."

Uzogara, Ekeoma E. 2019. "Dark and Sick, Light and Healthy: Black Women's Complexion-based Health Disparities." *Ethnicity & Health*, 24(2): 125–146.

Villarreal, Andrés. 2010. "Stratification by Skin Color in Contemporary Mexico." *American Sociological Review*, 75(5): 652–678.

Wade, T. Joel, Melanie Judkins Romano, and Leslie Blue. 2004. "The Effect of African American Skin Color on Hiring Preferences." *Journal of Applied Social Psychology*, 34(12): 2550–2558.

Washington, Scott Leon. 2011. "Hypodescent: A History of the Crystallization of the One-Drop Rule in the United States, 1880-1940." PhD diss. Princeton University.

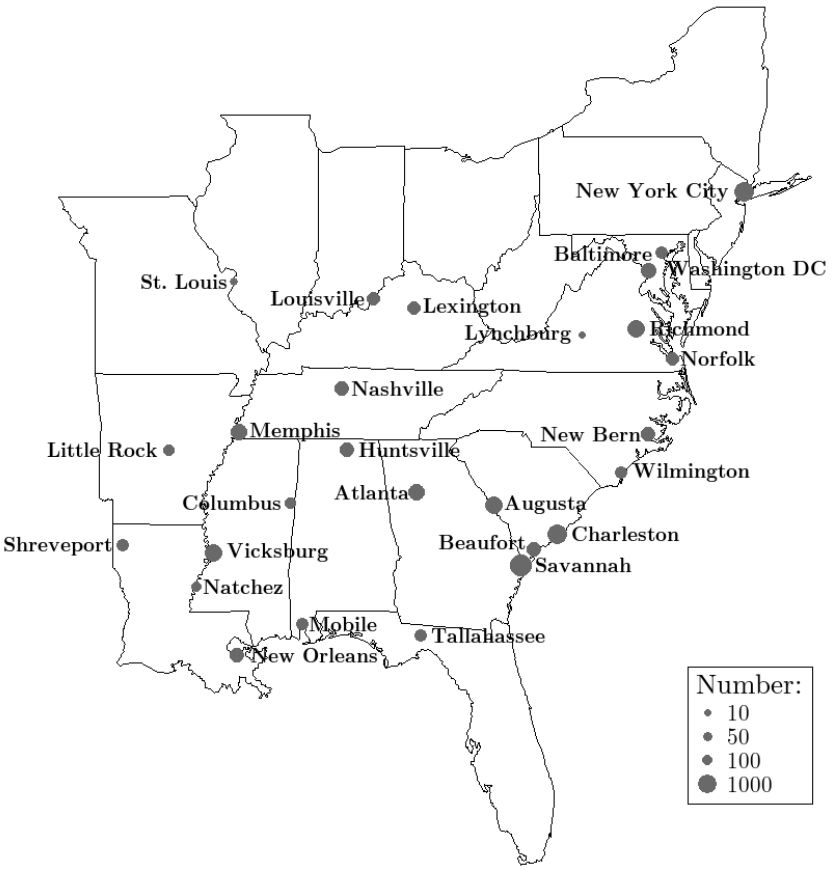
Waters, Mary C. 1994. "Ethnic and Racial Identities of Second-Generation Black Immigrants in New York City." *International Migration Review*, 28(4): 795–820.

Witzig, Ritchie. 1996. "The Medicalization of Race: Scientific Legitimization of a Flawed Social Construct." *Annals of Internal Medicine*, 125(8): 675–679.

Woo-Mora, L. Guillermo. 2024. "Unveiling the Cosmic Race: Intergenerational Skin Tone Gaps in Latin America and the Caribbean." *Working Paper*.

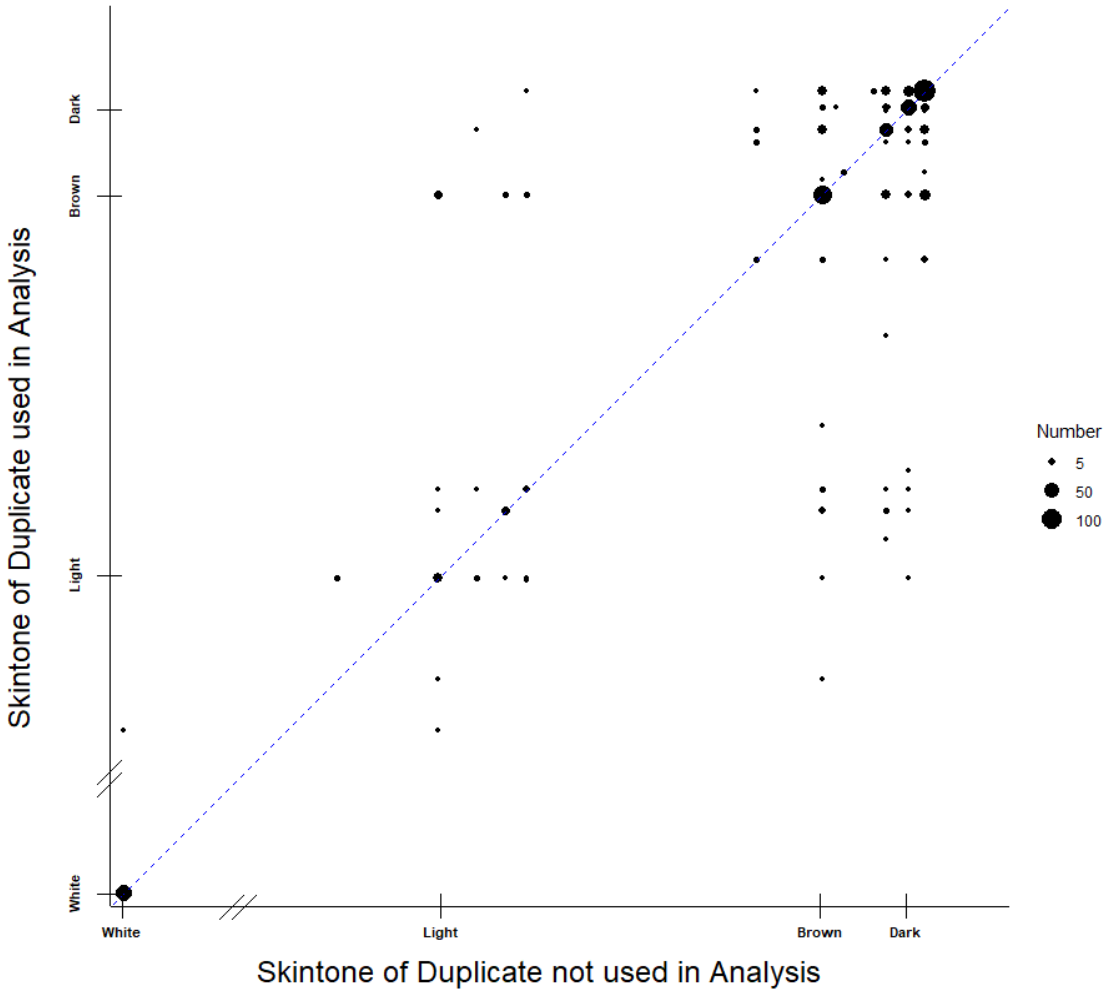
Zuberi, Tukufu. 2001. *Thicker Than Blood: How Racial Statistics Lie*. University of Minnesota Press.

Figure 1. Freedman Bank Depositor Accounts Linked to 1870 Census, by Branch



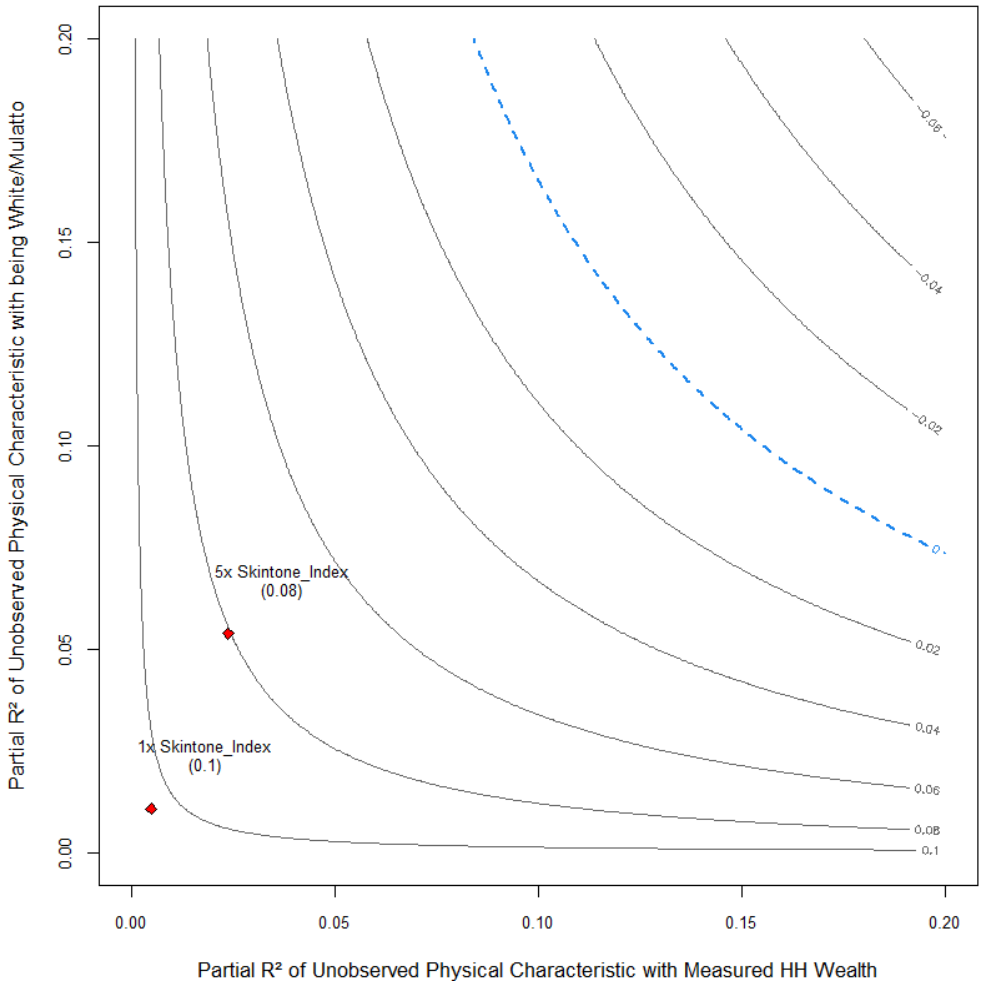
Notes: This Figure shows the locations of Freedman Bank branches in our sample, where the size of the circle is proportional to the number of depositors we link to the Census.

Figure 2. Comparing Repeated Bank Records on Depositor Skin Tone



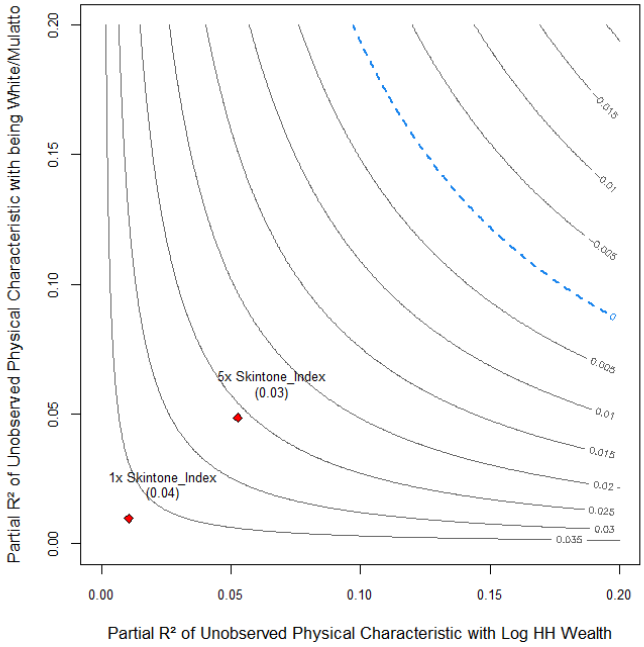
Notes: For 650 Freedman Bank depositors that opened an account twice and thereby have two measures of skin tone, this Figure plots the two skin tone index values against each other: the y-axis has the value used in the main analysis (from the depositor record with more biographical information) and the x-axis has the value not used in the main analysis (which we use as an instrument in some specifications). Each axis is labeled with reference values for the skin tone index (corresponding to white, light, brown, and dark in Table 1). Note the condensed axis between white and light. The size of the circle is proportional to the number of observations, and these circles are plotted against a dashed 45-degree line.

Figure 3. Bounding the Influence of Unobserved Physical Characteristics: Effect of Measured Wealth on Racialization as White or Mulatto (vs. Black)

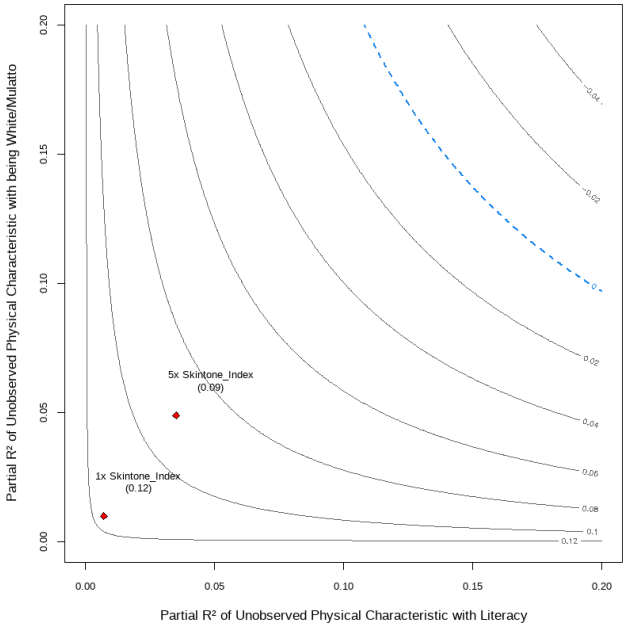


Notes: This Figure shows the potential influence of unobserved physical characteristics on the estimated effect of measured wealth on people being racialized in the 1870 Census as White or Mulatto (as opposed to Black). The first red diamond, closest to the origin, indicates the estimated effect of wealth if there were unobserved physical characteristics that were as jointly predictive of measured wealth and racialization as the detailed skin tone index (beyond three coarse skin tone groups). The second red diamond corresponds to unobserved physical characteristics that are 5 times as influential. To fully explain the effect, reaching the dashed blue line, unobserved physical characteristics would need to be 19 times as jointly predictive of measured wealth and racialization as the detailed skin tone index. If unobserved physical characteristics explained all residual variation in race, moving up the y-axis to one, then these unobserved characteristics would need to be 3.6 times more correlated with having measured wealth than the detailed skin tone index.

Figure 4. Bounding the Influence of Unobserved Physical Characteristics:
Panel A. Effect of Log Wealth on Racialization as White or Mulatto (vs. Black)

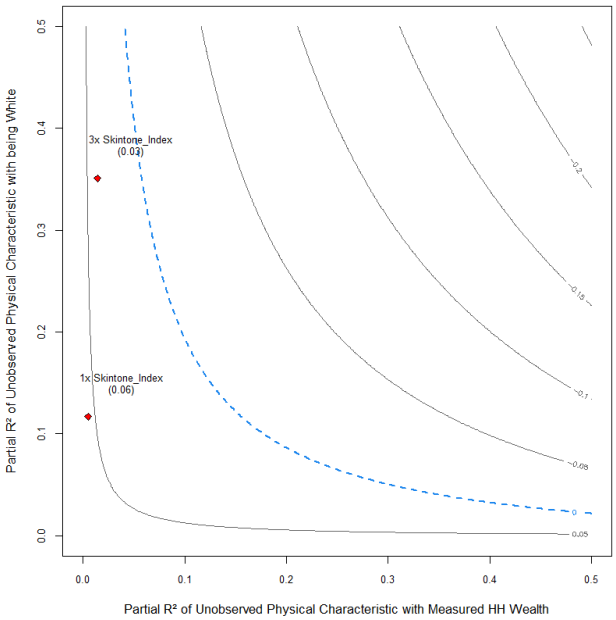


Panel B. Effect of Literacy on Racialization as White or Mulatto (vs. Black)

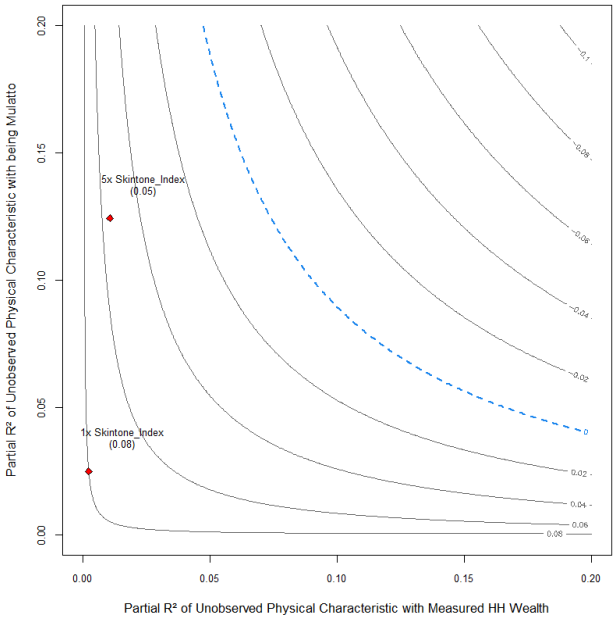


Notes: Similar to Figure 3, this Figure shows the potential influence of unobserved physical characteristics on the estimated effect of log wealth (Panel A) and literacy (Panel B) on people being racialized in the 1870 Census as White or Mulatto (as opposed to Black). For Panel A, to fully explain the effect of log wealth, unobserved physical characteristics would need to be 14 times as jointly predictive of log wealth and racialization as the detailed skin tone index. If unobserved physical characteristics explained all residual variation in race, moving up the y-axis to one, then these unobserved physical characteristics would need to be 1.9 times more correlated with log wealth than the detailed skin tone index. For Panel B, to fully explain the effect of literacy, unobserved physical characteristics would need to be 19 times as jointly predictive of literacy and racialization as the detailed skin tone index. If these characteristics fully explained all residual variation in race, they would need to be 4.0 times more correlated with literacy than the detailed skin tone index.

Figure 5. Bounding the Influence of Unobserved Physical Characteristics:
Panel A. Effect of Measured Wealth on Racialization as White (vs. Mulatto or Black)



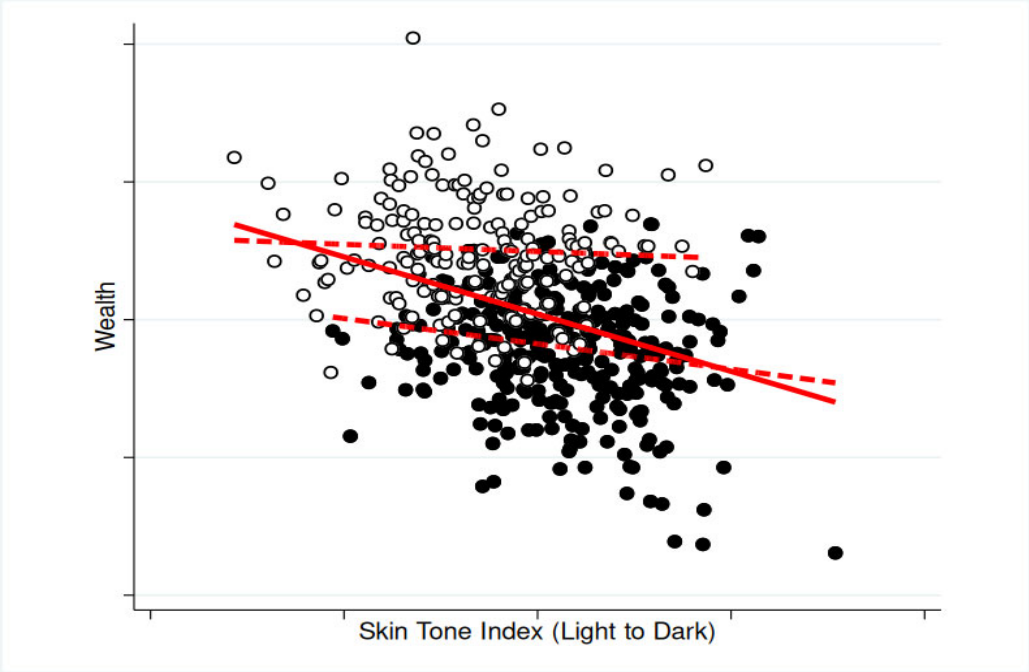
Panel B. Effect of Measured Wealth on Racialization as Mulatto (vs. Black)



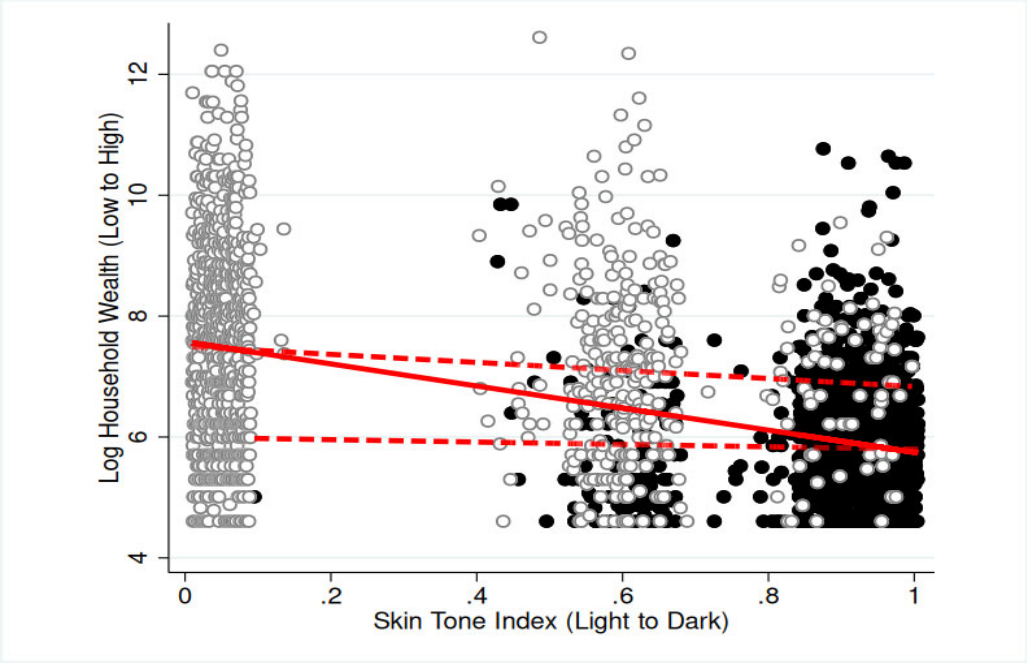
Notes: Similar to Figure 3, this Figure shows the potential influence of unobserved physical characteristics on the estimated effect of having measured wealth on people being racialized in the 1870 Census as White as opposed to Mulatto or Black (in Panel A) and being racialized as Mulatto as opposed to Black (in Panel B). For Panel A, to fully explain the effect of having measured wealth, unobserved physical characteristics would need to be 7 times as jointly predictive of measured wealth and racialization as the detailed skin tone index. If unobserved physical characteristics explained all residual variation in race, moving up the y-axis to one, then these unobserved physical characteristics would need to be 4.2 times more correlated with having measured wealth than the detailed skin tone index. For Panel B, to fully explain the effect of having measured wealth, unobserved physical characteristics would need to be 14 times as jointly predictive of having measured wealth and racialization as the detailed skin tone index. If these characteristics fully explained all residual variation in race, they would need to be 5.0 times more correlated with having measured wealth than the detailed skin tone index.

Figure 6: Simulated and Empirical Correlations Between Wealth and Skin Tone

Panel A: Simulated Correlations, Unconditional and Conditional on Census Race



Panel B: Empirical Correlations, Unconditional and Conditional on Census Race



Notes: Panel A shows simulated correlations between wealth and skin tone across all people (solid red line) and the attenuated relationships within each selected racial group (two dashed lines). Panel B shows empirical correlations between wealth and skin tone, across all people with measured wealth (solid red line) and within the groups racialized as White or Black in the 1870 Census (two dashed lines). In both Panels, the hollow circles are people racialized as White and the black circles are people racialized as Black.

Table 1. Freedman Bank Depositor Complexion and 1870 Census Race

Depositor Complexion	Number of Depositors	By 1870 Census Race:			Skin Tone Index
		Black	Mulatto	White	
	(1)	(2)	(3)	(4)	(5)
Black	3,606	3,383	198	25	0.97
Colored	95	86	9	0	0.95
Dark	1,633	1,489	131	13	0.95
Very Dark	49	44	5	0	0.95
Dark Black	31	27	4	0	0.94
Dark Brown	1,555	1,366	173	16	0.93
Very Dark Brown	72	61	11	0	0.92
Medium Dark Brown	19	17	1	1	0.92
Medium Black	89	72	16	1	0.90
Medium Brown	98	80	15	3	0.89
Brown	2,967	2,290	650	27	0.88
Nearly Black	183	145	13	25	0.83
Light Black	32	18	13	1	0.77
Dark Mulatto	20	10	9	1	0.73
Very Light Brown	13	6	6	1	0.69
Light Colored	6	2	4	0	0.67
Griff	45	27	5	13	0.66
Yellow	441	192	181	68	0.64
Light Brown	532	249	165	118	0.62
Mulatto	135	56	50	29	0.60
Light	715	273	266	176	0.57
Bright	83	30	34	19	0.57
Very Light	67	18	29	20	0.49
Nearly White	53	11	25	17	0.44
Olive	8	3	1	4	0.44
Fair	15	1	1	13	0.10
White	1,576	8	20	1,548	0.01
Total	14,138	9,964	2,035	2,139	0.78

Notes: Column 1 reports the total number of Freedman Bank depositors with the indicated complexion in the Bank's depositor record. We use these 27 detailed skin tone categories, which we standardize from 368 unique strings that include spelling variations and occasional other physical descriptives. Columns 2, 3, and 4 report the number of depositors racialized as Black, Mulatto, or White in the 1870 Census. Column 5 reports a continuous ``Skin Tone Index" that is defined for each skin tone category as: the share of depositors racialized as Black plus one-half times the share of depositors racialized as Mulatto.

Table 2. Depositor Summary Statistics in 1870 Census

	Sample	Differences from Full Census:		
	Mean	Black	Mulatto	White
	(1)	(2)	(3)	(4)
Adult	0.73	-	-	-
	[0.45]			
Male	0.72	-	-	-
	[0.45]			
Black	0.70	-	-	-
	[0.46]			
Mulatto	0.14	-	-	-
	[0.35]			
Measured Household Wealth	0.27	0.041	0.091	-0.006
	[0.44]	(0.004)	(0.010)	(0.009)
Household Wealth	$769	$3.3	$86.5	-$39.7
	[6590]	(37.0)	(134.5)	(473.6)
Literate	0.35	0.021	0.079	0.011
	[0.48]	(0.003)	(0.010)	(0.008)
Number of Depositors	14,138	9,964	2,035	2,139

Notes: Column 1 reports average characteristics for our sample of Freedman Bank depositors linked to the 1870 Census. Columns 2, 3, and 4 report differences for depositors by Census race, relative to non-depositors in the full Census, controlling for: county fixed effects and indicator variables for adult (18+) and male. We exclude all non-depositors in households with depositors. Literate is defined for people ages 10 and older. Standard deviations are reported in brackets, and robust standard errors are reported in parantheses.

Table 3. Estimated Differences in Census Racial Assignment by Socio-Economic Status

	White or Mulatto vs. Black		White vs. Mulatto or Black		Mulatto vs. Black	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Measured Wealth						
Wealth ≥ $100	0.109	0.108	0.068	0.068	0.087	0.085
	(0.008)	(0.007)	(0.005)	(0.005)	(0.009)	(0.009)
Skin Tone Index	-0.289		-0.275		-0.307	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999
Panel B: Log Wealth						
Log Wealth	0.040	0.041	0.031	0.030	0.044	0.043
	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)
Skin Tone Index	-0.281		-0.267		-0.302	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999
Panel C: Literacy						
Literate	0.124	0.122	0.065	0.067	0.094	0.092
	(0.008)	(0.008)	(0.005)	(0.005)	(0.008)	(0.008)
Skin Tone Index	-0.283		-0.271		-0.301	
	(0.004)		(0.003)		(0.011)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	13,029	13,029	13,029	13,029	11,277	11,277

Notes: Panel A, Columns 1 and 2, report effects of the household having measured wealth (≥ $100) on whether depositors are racialized in the 1870 Census as White or Mulatto (as opposed to Black). Columns 3 and 4 report effects on whether depositors are racialized as White (as opposed to Mulatto or Black), and Columns 5 and 6 report effects on whether depositors are racialized as Mulatto (as opposed to Black, excluding White depositors from the sample). Panel B reports effects of the household's log wealth, adding $45.25 to all households with unmeasured wealth. Panel C reports effects of the depositor being literate, which is limited to people ages 10 and older in the Census. Columns 1, 3, and 5 control for the skin tone index (from Column 5 of Table 1), which we standardize to have a mean of zero and a standard deviation of one. Columns 2, 4, and 6 control for fixed effects for each of the 27 detailed skin tone categories (listed in Table 1). All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Table 4. Effects of Socio-Economic Status on Census Racial Assignment, by Skin Tone Group

	Skin Tone Group:				
	Black	Dark	Medium	Light	White
	(1)	(2)	(3)	(4)	(5)
Panel A: White or Mulatto vs. Black					
Wealth ≥ $100	0.058	0.085	0.146	0.222	0.010
	(0.013)	(0.016)	(0.018)	(0.021)	(0.004)
Log Wealth	0.032	0.040	0.067	0.083	0.002
	(0.007)	(0.007)	(0.007)	(0.005)	(0.001)
Literate	0.034	0.067	0.144	0.306	0.035
	(0.011)	(0.014)	(0.018)	(0.022)	(0.019)
Number of Depositors	3,701	3,359	3,389	2,113	1,576
Panel B: White vs. Mulatto or Black					
Wealth ≥ $100	0.014	0.033	0.049	0.250	0.012
	(0.005)	(0.008)	(0.008)	(0.021)	(0.007)
Log Wealth	0.012	0.017	0.026	0.093	0.003
	(0.004)	(0.004)	(0.004)	(0.006)	(0.002)
Literate	0.014	0.022	0.031	0.256	0.067
	(0.005)	(0.005)	(0.007)	(0.020)	(0.026)
Number of Depositors	3,701	3,359	3,389	2,113	1,576
Panel C: Mulatto vs. Black					
Wealth ≥ $100	0.046	0.057	0.113	0.128	
	(0.012)	(0.015)	(0.017)	(0.028)	
Log Wealth	0.022	0.026	0.051	0.067	
	(0.006)	(0.007)	(0.007)	(0.009)	
Literate	0.022	0.049	0.124	0.229	
	(0.010)	(0.013)	(0.018)	(0.027)	
Number of Depositors	3,676	3,329	3,331	1,635	28

Notes: Each reported coefficient is from a separate regression, as in Table 3 (Columns 2, 4, and 6 of Panels A, B, and C). Each column splits the analysis from Table 3 into separate skin tone groups. These skin tone groups are defined from the detailed skin tone categories in Table 1: Black (black and colored); Dark (dark to medium dark brown); Medium (medium black to dark mulatto); Light (very light brown to fair); and White (white). The samples are moderately smaller for the Literate row, as this question is limited to people ages 10 and older in the Census. All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Table 5. Geographic Heterogeneity in Census Racial Assignment by Socio-economic Status

	Baseline	One Standard Deviation Difference in County Characteristics:		
	Average Effect	Lower Black Wealth	Lower Black Literacy	More Racial Violence
	(1)	(2)	(3)	(4)
Panel A: White or Mulatto vs. Black				
Wealth ≥ $100	0.108	0.011	0.018	0.028
	(0.007)	(0.006)	(0.005)	(0.010)
Log Wealth	0.041	0.008	0.012	0.015
	(0.002)	(0.001)	(0.001)	(0.004)
Literate	0.122	-0.009	0.040	0.002
	(0.008)	(0.005)	(0.007)	(0.010)
Number of Depositors	14,138	14,124	14,124	8,060
Panel B: White vs. Mulatto or Black				
Wealth ≥ $100	0.068	0.008	0.009	0.031
	(0.005)	(0.005)	(0.004)	(0.008)
Log Wealth	0.030	0.007	0.008	0.019
	(0.002)	(0.001)	(0.001)	(0.003)
Literate	0.067	-0.018	0.013	0.021
	(0.005)	(0.004)	(0.004)	(0.008)
Number of Depositors	14,138	14,124	14,124	8,060
Panel C: Mulatto vs. Black				
Wealth ≥ $100	0.085	-0.016	-0.002	0.014
	(0.009)	(0.012)	(0.008)	(0.012)
Log Wealth	0.043	-0.006	0.000	0.011
	(0.004)	(0.004)	(0.004)	(0.006)
Literate	0.092	-0.007	0.034	-0.013
	(0.008)	(0.010)	(0.007)	(0.010)
Number of Depositors	11,999	11,999	11,999	7,566

Notes: Column 1 reports our baseline average effects from Table 3. Columns 2, 3, and 4 report how this effect varies in counties with one standard deviation lower Black wealth (Column 2), lower Black literacy (Column 3), and more racial violence (Column 4). These estimates come from separate regressions, where we estimate the main effect and an interaction with the indicated county characteristic. For Column 2, we use the county's average wealth of people racialized as Black in the 1870 Census: total wealth of Black people in the county, divided by the number of Black people in the county. For Column 3, we use the county's literacy rate (for people ages 10 and older) among people racialized as Black in the 1870 Census. For Columns 2 and 3, we exclude depositors and their family members in calculating county averages, and the sample declines slightly because of 12 depositors in counties with zero Black people in the Census. For Column 4, we use data from Chyn, Haggag, and Stuart (2024) on the frequency of crimes reported by the Freedmen's Bureau for racial violence perpetrated against African Americans. For each county, we calculate the number of crimes reported per person racialized as Black or Mulatto, conditional on the county having at least one report of violence. The sample declines because of non-reported data for some states. All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Table 6. Differences in Census Racial Assignment by Bank Record Characteristics

	White or Mulatto vs. Black		White vs. Mulatto or Black		Mulatto vs. Black	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Distinctively Black Name						
Distinctively Black Name	-0.021	-0.019	-0.007	-0.009	-0.014	-0.013
	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)
Skin Tone Index	-0.299		-0.283		-0.310	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,057	14,057	14,057	14,057	11,926	11,926
Panel B: Occupation in Bank Record						
Professional Occupation	0.057	0.052	0.013	0.020	0.058	0.051
	(0.012)	(0.012)	(0.008)	(0.007)	(0.015)	(0.015)
Service Occupation	0.042	0.030	-0.008	0.005	0.034	0.028
	(0.008)	(0.008)	(0.004)	(0.004)	(0.008)	(0.008)
Skin Tone Index	-0.305		-0.282		-0.302	
	(0.004)		(0.003)		(0.012)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	10,446	10,446	10,446	10,446	9,267	9,267
Panel C: Writes Own Signature in Bank Record						
Depositor Signature	0.055	0.051	0.000	0.004	0.059	0.054
	(0.008)	(0.008)	(0.004)	(0.004)	(0.008)	(0.008)
Skin Tone Index	-0.295		-0.284		-0.318	
	(0.004)		(0.003)		(0.011)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	12,110	12,110	12,110	12,110	10,334	10,334

Notes: This Table reports estimates similar to those in Table 3, but for depositor characteristics defined using the Bank's depositor record. Panel A reports effects of having a one standard deviation more distinctively Black name, based on the frequency of that name in the 1870 Census among people racialized as Black or Mulatto compared to those racialized as White. Panel B reports effects of having a professional occupation or service occupation, relative to the omitted category of having a farming or laborer occupation. Pancel C reports effects associated with the depositor having signed their name in the Bank record (as opposed to the Bank cashier signing the name and indicating ``his mark.") All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Table 7. Differences in Wealth and Literacy by Skin Tone, Attenuated Within Census Race

			Within Census Race:		
	Full Sample		Black	Mulatto	White
	(1)	(2)	(3)	(4)	(5)
Panel A: Wealth ≥ $100					
Skin Tone Index	-0.122	-0.011	-0.036	-0.034	0.009
	(0.005)	(0.008)	(0.012)	(0.018)	(0.011)
Black		-0.383			
		(0.021)			
Mulatto		-0.258			
		(0.021)			
Number of Depositors	14,138	14,138	9,964	2,035	2,139
Panel B: Log Wealth					
Skin Tone Index	-0.532	-0.107	-0.085	-0.226	-0.080
	(0.019)	(0.027)	(0.028)	(0.058)	(0.044)
Black		-1.476			
		(0.077)			
Mulatto		-1.119			
		(0.076)			
Number of Depositors	14,138	14,138	9,964	2,035	2,139
Panel C: Literate					
Skin Tone Index	-0.171	-0.055	-0.061	-0.131	-0.076
	(0.005)	(0.008)	(0.014)	(0.020)	(0.010)
Black		-0.399			
		(0.021)			
Mulatto		-0.246			
		(0.021)			
Number of Depositors	13,029	13,029	9,448	1,829	1,752

Notes: Column 1 reports the unconditional relationship between the indicated outcome (in each Panel) and the skin tone index. Column 2 also includes whether the person is racialized as Black or racialized as Mulatto. Columns 3, 4, and 5 report separate estimates within the subsamples of depositors who are racialized as Black, Mulatto, or White. All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Table 8. Differences in Wealth and Literacy by Skin Tone, 2SLS with Multiple Measures

	OLS		2SLS	
	(1)	(2)	(3)	(4)
Panel A: Wealth ≥ $100				
Skin Tone Index	-0.121	-0.038	-0.114	-0.010
	(0.025)	(0.044)	(0.028)	(0.053)
Black		-0.284		-0.345
		(0.134)		(0.147)
Mulatto		-0.150		-0.201
		(0.133)		(0.142)
Number of Depositors	650	650	650	650
Panel B: Log Wealth				
Skin Tone Index	-0.485	-0.213	-0.510	-0.219
	(0.088)	(0.138)	(0.099)	(0.160)
Black		-0.993		-0.979
		(0.409)		(0.423)
Mulatto		-0.670		-0.658
		(0.391)		(0.399)
Number of Depositors	650	650	650	650
Panel C: Literate				
Skin Tone Index	-0.173	-0.129	-0.183	-0.145
	(0.029)	(0.046)	(0.032)	(0.056)
Black		-0.147		-0.114
		(0.129)		(0.146)
Mulatto		-0.002		0.026
		(0.127)		(0.139)
Number of Depositors	612	612	612	612
First Stage:				
Duplicate Skin Tone Index			0.926	0.776
			(0.023)	(0.037)
F-Statistic on Instrument			1626.049	438.694

Notes: Columns 1 and 2 replicate Table 7, but restricting the sample to depositors whose skin tone is recorded twice from opening multiple accounts with the Bank. Columns 3 and 4 are two-stage least squares estimates corresponding to Columns 1 and 2, but instrumenting for the primary value of the skin tone index (from the record with more biographical information) using the secondary value of the skin tone index. The first-stage coefficients are reported below, along with the F-statistic of the excluded instrument. All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Appendix for:

The Social Construction of Race after Emancipation: US Census Racial Assignment Based on Skin Tone and Socioeconomic Status

Anjali Adukia
University of Chicago

Richard Hornbeck
University of Chicago

Daniel Keniston
Louisiana State University

Benjamin Lualdi
University of Chicago

January 2025

A Linking Freedman Bank Depositor Records to the 1870 Census: Iterative Hand-Linking and Machine Learning Model Details

This section provides further details on our process for linking the Freedman Bank depositor records to the 1870 Census. In particular, we include details on the variables used in estimating the random forest linking model. We use the model to iterate on the hand-links, but we ultimately hand-link all records.

To begin the process, in a "blocking" stage, we generate up to 20 candidate matches in the 1870 Census for each of the 44,077 linkable depositors in the Freedman Bank depositor records. A linkable record has the depositor name, age, and birth place (state or country).

We separate the 44,077 linkable depositors into two groups: 35,958 adults and 8,119 children. We define a depositor as a child when their age in the depositor record is younger than 18 and they have no spouse or children in their depositor record.This distinction between adults and children determines the family information shown when hand-linking the records: for children, we display the names of their parents and siblings; for adults, we display the names of their spouse and children.

The blocking code provides 861,813 "potential matches" for the 44,077 depositors, for an average of 19.55 potential matches per linkable depositor.

We send these potential matches to our data-linking team, who use their judgment to determine the appropriate links given the provided information (depositor name, age, complexion/race, and family member names). Each depositor was hand-linked by two members of our data-linking team, with a third person resolving any differences in linking decisions.

The data-linking team determines for each "potential match" one of three categories of link:

- **Type 0**: If there is no likely match, then the team member enters "0" in all the rows corresponding to this depositor.
- **Type 1**: If there is one likely match, then the team member enters "1" in the row for that matched person in the 1870 Census and enters "0" for each other potential match.
- **Type 3**: If there are multiple candidates that could be a possible match, then our team member enters "3" in those rows and enters "0" elsewhere. For this paper, we exclude these one-to-many matches and focus only on the unique matches, though it is relevant that they had the option of indicating multiple matches.

After hand-linking all 44,077 depositor records, we construct two random forest models (one for adults and one for children) to help assess the potential quality of links made.

We use these assessments to revisit potentially bad matches (e.g., a Type 1 that should be Type 0) and revisit potentially missed matches (e.g., a Type 0 that should be a Type 1), as described in the main text. We also report robustness to excluding hand-links below particular predicted thresholds after re-estimating the model with the updated hand-links.

A.1 Defining Machine Learning Model Variables

For each random forest model, the dependent variable ("true match probability") is defined as: 0 for potential matches that are assigned a "0"; 1 for potential matches that are assigned a "1"; and $1/K$ for each potential match assigned a "3," where the depositor had $K \geq 2$ type 3 matches with the 1870 Census IDs.

We then create predictors for the random forest models, based flexible on the characteristics of the people in the depositor records and the 1870 Census. The random forest model for adults contains 40 predictors and the model for children has 46 predictors, where the model can iteratively determine which predictors to emphasize. For each pair of people (one from the depositor records and the other from the 1870 Census), the predictors are as follows:

Depositor String Similarity Variables

Variable(s)	Description
jw_first, jw_last	One minus the Jaro-Winkler distances between the first/last names of the depositor and potential Census match.
L1_first, L1_last	An indicator equal to one if the first letter of the first/last names are the same
inits12_first	An indicator equal to one if either of two first names only has one letter.
inits2_first	An indicator equal to one if both of two first names only have one letter.
namefrst_nickname_jw	One minus the smallest Jaro-Winkler distance between a first name and the other first name's corresponding nickname.
first_common, last_common	A commonness score of the depositor's first/last name.
nys_first, nys_last	Indicators equal to one if the phonetic spelling between the first/last names is the same.

Depositor Age and Complexion/Race

Variable(s)	Description
birthyear_diff	Difference in birth year.
race_1_blk, race_1_wht	Indicators equal to one if the person in the depositor records has darker than "very light brown" complexion in the depositor records or whether the person has "white" complexion, where the omitted category is what we define as ambiguous complexion (between "very light brown" and "fair").
race_wht_blk	A dummy variable equal to 1 if one person has complexion darker than "very light brown" in the depositor records (from Table 1) and is White in the Census, or if one person has "white" complexion in the depositor records and is Black or Mulatto in the Census.

Spouse/Parents String Similarity Variables

Variable(s)	Description
L1_spouse, L1_mother, L1_father	An indicator equal to one if the initial letters of the first/last names or the initial letters of the spouse or parents' names are the same.
jw_spouse, jw_mother, jw_father	One minus the Jaro-Winkler distances between the spouses'/parents' names.
jw_parents_mean	The average of jw_mother and jw_father.
spouse_common, mother_common, father_common	A commonness score of the depositor's spouse's/parents' first names.
nys_spouse, nys_mother, nys_father	Indicators equal to one if the phonetic spelling between the spouses'/parents' names is the same.
spouse_nickname_jw, mother_nickname_jw, father_nickname_jw	One minus the Jaro-Winkler distance between a spouse's/parent's name and the other spouse's/parent's corresponding nickname.

Spouse/Parents String Similarity Missing Variables

Variable(s)	Description
missing12_spouse, missing12_mother, missing12_father	An indicator equal to one if spouse/parent names are missing in either the depositor record or census record.
missing2_spouse, missing12_mother, missing12_father	An indicator equal to one if spouse/parent names are missing in both the depositor record and census record.

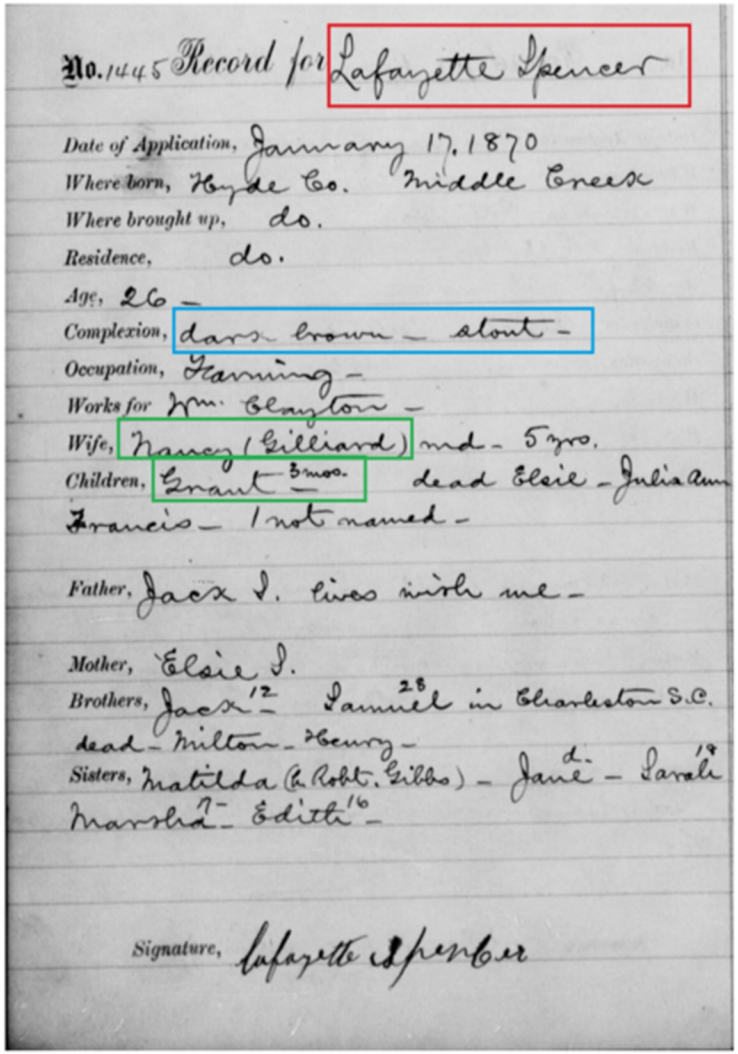
Children/Siblings String Similarity Variables

Variable(s)	Description
kid_nickname_jw_max, sibling_nickname_jw_max	One minus the minimum Jaro-Winkler distance between a child/sibling's name and another child/sibling's corresponding nickname.
jw_kids_max, jw_kids_mean, jw_siblings_max, jw_siblings_mean	One minus the minimum/mean of the Jaro-Winkler distances between the children/siblings' names.
jw_rsc_kids_max, jw_rsc_kids_mean, jw_rsc_siblings_max, jw_rsc_siblings_mean	One minus the minimum/mean of the Jaro-Winkler distances between the children/siblings' names, re-scaled by the kids' names' commonness.
nys_kids_any, nys_kids_mean, nys_siblings_any, nys_siblings_mean	Minimum/mean of the phonetic distances between the children/siblings' names.
nys_rsc_kids_any, nys_rsc_kids_mean, nys_rsc_siblings_any, nys_rsc_siblings_mean	Minimum/mean of the phonetic distances between the children/siblings' names, re-scaled by the children/siblings' names' commonness.

Children/Siblings String Similarity Missing Variables

Variable(s)	Description
missing12_kids, missing12_siblings	An indicator equal to one if children/siblings names are missing in either the depositor record or census record.
missing2_kids, missing12_siblings	An indicator equal to one if children/siblings names are missing in both the depositor record and census record.
nys_kids_any_missing, nys_kids_mean_missing, nys_siblings_any_missing, nys_siblings_mean_missing	Indicator equal to one if the corresponding variable is missing.
nys_rsc_kids_any_missing, nys_rsc_kids_mean_missing, nys_rsc_siblings_any_missing, nys_rsc_siblings_mean_missing	Indicator equal to one if the corresponding variable is missing.
jw_kids_max_missing, jw_siblings_max_missing, jw_rsc_kids_max_missing, jw_rsc_siblings_max_missing	Indicator equal to one if the corresponding variable is missing.

Appendix Figure 1. Freedman Bank Depositor Record for Lafeyette Spencer



No. 1445 Record for Lafayette Spencer

Date of Application, January 17, 1870
Where born, Hyde Co. Middle Creek
Where brought up, do.
Residence, do.
Age, 26 –
Complexion, dark brown – stout –
Occupation, Farming –
Works for Wm. Clayton –
Wife, Nancy (Gilliard) md – 5 yrs.
Children, Grant 3 mos. dead Elsie – Julia Ann
Francis – 1 not named –

Father, Jack S. lives with me –

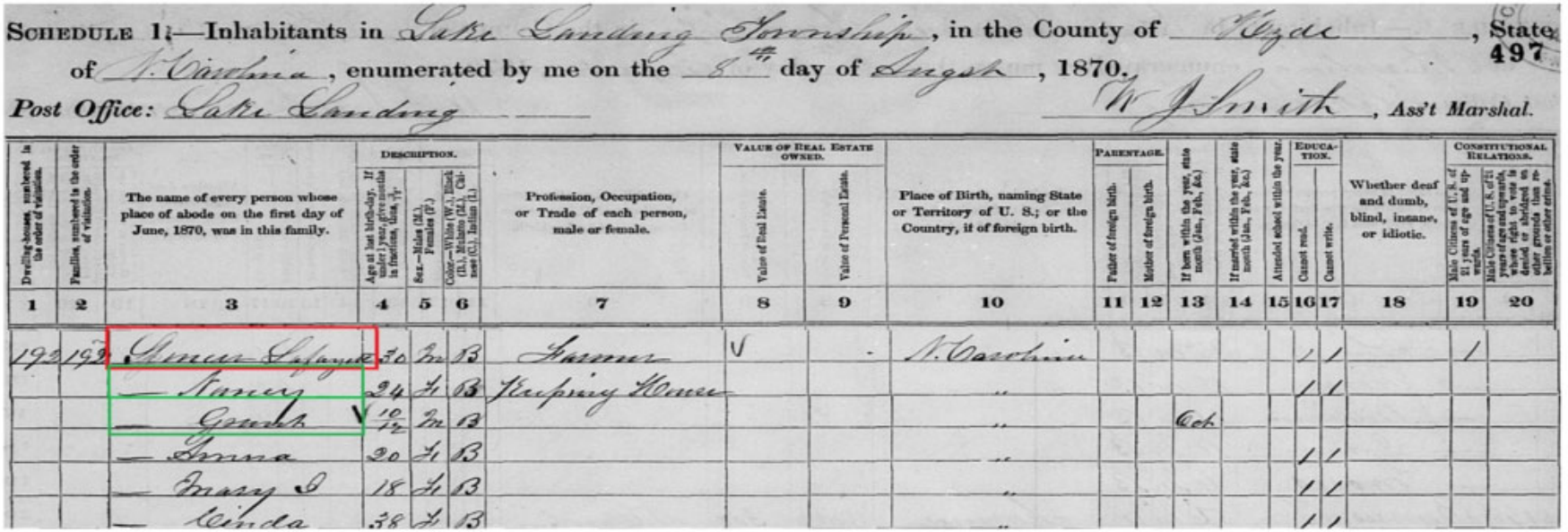
Mother, Elsie S.
Brothers, Jack 12 – Samuel 28 in Charleston S.C.
dead – Milton – Henry –
Sisters, Matilda (m. Robt. Gibbs) – Jane d. – Sarah 18
Marsha 7 – Edith 16 –

Signature, Lafayette Spencer

No. 1445 Record for	Lafayette Spencer
Date of Application	January 17, 1870
Where born	Hyde County, Middle Creek
Where brought up	"
Residence	"
Age	26 years old
Complexion	Dark brown, stout
Occupation	Farming
Works for	William Clayton
Wife	Nancy (Gilliard), married 5 yrs
Children	Grant 3 months; dead: Elsie, Julia Ann,
	Francis, 1 not named
Father	Jack S. lives with me
Mother	Elsie S.
Brothers	Jack 12, Samuel 28 in Charleston, S.C.
	dead: Milton, Henry
Sisters	Matilda (m. Robert Gibbs), Jane (d.), Sarah (18)
	Martha (17), Edith (16)

Notes: The image on the left is the depositor record for Lafayette Spencer, from an account opened on January 17, 1870 with the New Bern branch of the Freedman Bank. On the right is a typed transcription. Note that Hyde County is near New Bern (Craven County), but is not a neighboring county.

Appendix Figure 2. 1870 Census Record for Lafayette Spencer



SCHEDULE 1.—Inhabitants in Lake Landing Township, in the County of Hyde, State of N. Carolina, enumerated by me on the 8th day of August, 1870. 497

Post Office: Lake Landing — W. J. Smith, Ass't Marshal.

Dwelling-houses, numbered in the order of visitation.	Families, numbered in the order of visitation.	The name of every person whose place of abode on the first day of June, 1870, was in this family.	Description: Age at last birth-day. If under 1 year, give months in fractions, thus, 7/12.	Description: Sex.—Males (M.), Females (F.)	Description: Color—White (W.), Black (B.), Mulatto (M.), Chinese (C.), Indian (I.)	Profession, Occupation, or Trade of each person, male or female.	Value of Real Estate owned.	Value of Personal Estate.	Place of Birth, naming State or Territory of U. S.; or the Country, if of foreign birth.	Parentage: Father of foreign birth.	Parentage: Mother of foreign birth.	If born within the year, state month (Jan., Feb., &c.)	If married within the year, state month (Jan., Feb., &c.)	Attended school within the year.	Education: Cannot read.	Education: Cannot write.	Whether deaf and dumb, blind, insane, or idiotic.	Constitutional Relations: Male Citizens of U.S. of 21 years of age and upwards.	Constitutional Relations: Male Citizens of U.S. of 21 years of age and upwards, whose rights to vote is denied or abridged on other grounds than rebellion or other crime.
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
193	193	Spencer Lafayette	30	M	B	Farmer	0	0	N. Carolina							✓		✓	
		" Nancy	24	F	B	Keeping House	0	0	"							✓		✓	
		" Grant	0	M	B		0	0	"			Oct							
		" Arrna	20	F	B		0	0	"							✓		✓	
		" Mary L	10	F	B		0	0	"							✓		✓	
		" Cinda	38	F	B		0	0	"							✓		✓	

Notes: The image above is the 1870 Census record for Lafayette Spencer, with a typed transcription below.

Appendix Table 1. Tabulation of Bank Depositor Complexion Categories by Branch

	Freedman Bank Branch:																										
Complexion:	ATL	AUG	BAL	BEA	CHA	CMS	HUN	LEX	LIT	LOU	LYN	MEM	MOB	NAS	NAT	NOR	NWB	NWL	NYC	RIC	SAV	SHR	STL	TAL	VIC	WDC	WIL
Black	312	372	55	137	379	2	110	94	13	14	0	175	39	6	43	189	180	185	58	309	384	48	2	41	430	4	25
Bright	4	7	5	0	0	3	1	3	0	6	0	11	0	15	0	3	3	1	0	1	0	0	0	17	0	3	0
Brown	183	75	62	152	286	10	134	14	60	198	3	180	73	117	9	3	78	18	128	199	324	43	0	56	304	210	48
Colored	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	90	2	0
Dark	5	0	27	2	22	38	18	3	1	11	1	77	69	82	4	78	33	10	1	97	554	67	3	105	5	253	67
Dark Black	0	1	0	0	1	0	0	26	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Dark Brown	88	105	14	88	393	20	10	30	38	34	0	95	23	117	3	0	52	9	80	232	68	7	0	22	3	9	15
Dark Mulatto	0	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Fair	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	0
Griff	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	38	0	0	0	0	6	0	0	1	0
Light	4	22	1	6	139	18	2	12	2	2	0	52	8	5	2	11	36	14	35	4	247	30	0	8	13	28	14
Light Black	1	0	0	0	3	0	0	26	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Light Brown	30	12	2	9	47	19	18	8	20	55	0	39	3	42	7	1	4	2	77	70	12	11	0	16	6	12	10
Light Colored	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Medium Black	0	0	0	0	0	0	1	87	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Medium Brown	1	0	0	0	0	23	26	15	1	0	0	2	0	8	0	0	16	2	0	1	0	0	0	0	0	1	2
Med. Dark Brown	0	0	0	0	0	4	2	2	0	0	0	4	0	3	0	1	1	0	0	2	0	0	0	0	0	0	0
Mulatto	0	51	1	0	0	5	0	0	0	43	0	1	1	0	1	0	1	5	0	3	6	0	1	6	2	8	0
Nearly Black	55	108	0	0	0	0	0	4	0	0	0	4	0	0	0	0	2	0	0	8	2	0	0	0	0	0	0
Nearly White	11	9	0	1	6	0	3	1	2	0	0	2	0	0	0	4	2	0	5	4	2	0	0	0	0	1	0
Olive	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	1	0	0	0	0	0	1	0
Very Dark	7	0	0	0	1	3	0	1	2	0	0	1	0	4	0	0	1	0	0	16	0	0	1	0	0	0	12
Very Dark Brown	51	0	0	0	0	0	1	1	1	0	0	1	0	2	0	0	0	0	0	13	0	0	0	0	0	0	2
Very Light	1	0	0	0	10	0	0	0	5	0	0	1	0	2	0	2	2	1	12	7	21	0	0	1	0	1	1
Very Light Brown	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	1	1
White	53	26	59	5	55	0	14	19	47	16	0	6	0	10	10	14	41	65	944	11	76	5	0	6	51	26	17
Yellow	4	53	23	1	1	9	71	1	4	2	0	28	1	18	13	16	0	73	1	2	9	1	0	1	62	45	2

Notes: For each Freedman Bank branch, the columns list their number of depositors by detailed skin tone category. The branches are ordered as follows: Atlanta (GA), Augusta (GA), Baltimore (MD), Beaufort (SC), Charleston (SC), Columbus (MS), Hunstville (AL), Lexington (KY), Little Rock (AR), Louisville (KY), Lynchburg (VA), Memphis (TN), Mobile (AL), Nashville (TN), Natchez (MS), Norfolk (VA), New Bern (NC), New Orleans (LA), New York (NY), Richmond (VA), Savannah (GA), Shreveport (LA), St. Louis (MO), Tallahassee (FL), Vicksburg (MS), Washington (DC), Wilmington (NC).

Appendix Table 2. Effects of SES on Census Racial Assignment, Robustness to Link Restrictions

	Baseline Handlinks	Exclude ML Prob < 0.10	Exclude ML Prob < 0.50	≥ 1 Family Match	≥ 2 Family Match	Exact Name & Birth Year ±1
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: White or Mulatto vs. Black						
Wealth ≥ $100	0.108	0.099	0.077	0.101	0.093	0.100
	(0.007)	(0.008)	(0.009)	(0.009)	(0.012)	(0.014)
Log Wealth	0.041	0.037	0.029	0.038	0.033	0.032
	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
Literate	0.122	0.110	0.094	0.121	0.118	0.093
	(0.008)	(0.008)	(0.010)	(0.011)	(0.014)	(0.015)
Number of Depositors	14,138	12,459	8,598	8,335	4,974	3,568
Panel B: White vs. Mulatto or Black						
Wealth ≥ $100	0.068	0.056	0.028	0.062	0.047	0.047
	(0.005)	(0.005)	(0.005)	(0.006)	(0.007)	(0.009)
Log Wealth	0.030	0.025	0.013	0.026	0.021	0.018
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Literate	0.067	0.051	0.028	0.060	0.052	0.042
	(0.005)	(0.004)	(0.004)	(0.007)	(0.008)	(0.008)
Number of Depositors	14,138	12,459	8,598	8,335	4,974	3,568
Panel C: Mulatto vs. Black						
Wealth ≥ $100	0.085	0.081	0.074	0.079	0.083	0.099
	(0.009)	(0.009)	(0.011)	(0.011)	(0.014)	(0.018)
Log Wealth	0.043	0.041	0.036	0.041	0.039	0.048
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.008)
Literate	0.092	0.090	0.085	0.095	0.097	0.082
	(0.008)	(0.009)	(0.010)	(0.011)	(0.014)	(0.016)
Number of Depositors	11,999	10,769	7,561	6,957	4,025	2,888

Notes: Column 1 reports our baseline estimates from Table 3 (in Columns 2, 4, and 6 of Panels A, B, and C). Each coefficient is from a separate regression and, as before, the sample size is smaller for Literate. In Column 2 and 3, we exclude links that the machine-learning model assigns a probability below 0.10 or 0.50. In Column 4, we restrict the sample to depositors with at least one other household member whose name also matches between the depositor record and 1870 Census (e.g., a spouse or child for adult depositors, and a parent or sibling for child depositors). In Column 5, we restrict the sample to depositors with at least two other household members whose name matches. Within our hand-linked households, for this sample restriction we define a household member name "match" as having Jaro-Winkler string distance below 0.30 (searching within family members only). In Column 6, we limit the hand-linked records to those with an exact name match and an age difference that is one year or less. All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Appendix Table 3. Census Racial Assignment by Depositor Characteristics in Bank Records, Robustness to Link Restrictions

	Baseline Handlinks	Exclude ML Prob < 0.10	Exclude ML Prob < 0.50	≥ 1 Family Match	≥ 2 Family Match	Exact Name & Birth Year ±1
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: White or Mulatto vs. Black						
Blackness of Name	-0.019	-0.017	-0.015	-0.018	-0.019	-0.035
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)
Professional Occupation	0.052	0.048	0.053	0.061	0.061	0.097
	(0.012)	(0.013)	(0.015)	(0.017)	(0.022)	(0.024)
Service Occupation	0.030	0.033	0.039	0.044	0.057	0.026
	(0.008)	(0.008)	(0.010)	(0.011)	(0.015)	(0.016)
Own Signature	0.051	0.056	0.058	0.052	0.066	0.066
	(0.008)	(0.008)	(0.009)	(0.010)	(0.013)	(0.015)
Number of Depositors	14,138	12,459	8,598	8,335	4,974	3,568
Panel B: White vs, Mulatto or Black						
Blackness of Name	-0.009	-0.006	-0.002	-0.008	-0.009	-0.009
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Professional Occupation	0.020	0.016	0.013	0.023	0.029	0.042
	(0.007)	(0.007)	(0.007)	(0.010)	(0.013)	(0.015)
Service Occupation	0.005	0.008	0.008	0.007	0.008	-0.001
	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)	(0.008)
Own Signature	0.004	0.002	0.000	-0.001	0.004	0.013
	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)	(0.008)
Number of Depositors	14,138	12,459	8,598	8,335	4,974	3,568
Panel C: Mulatto vs. Black						
Blackness of Name	-0.013	-0.013	-0.014	-0.013	-0.013	-0.030
	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)
Professional Occupation	0.051	0.046	0.053	0.057	0.054	0.108
	(0.015)	(0.015)	(0.018)	(0.020)	(0.028)	(0.032)
Service Occupation	0.028	0.030	0.035	0.044	0.059	0.028
	(0.008)	(0.009)	(0.011)	(0.012)	(0.016)	(0.017)
Own Signature	0.054	0.059	0.062	0.061	0.076	0.067
	(0.008)	(0.008)	(0.010)	(0.011)	(0.014)	(0.016)
Number of Depositors	11,999	10,769	7,561	6,957	4,025	2,888

Notes: Column 1 reports our estimates from Table 6 (in Columns 2, 4, and 6 of Panels A, B, and C). Columns 2 to 6 report estimates for restricted samples, as described in Appendix Table 2 (that correspond to our baseline estimates from Table 3). In Columns 2 and 3, we exclude links that the machine-learning model assigns a probability below 0.10 or 0.50. In Column 4, we restrict the sample to depositors with at least one other household member whose name also matches between the depositor record and 1870 Census (e.g., a spouse or child for adult depositors, and a parent or sibling for child depositors). In Column 5, we restrict the sample to depositors with at least two other household members whose name matches. Within our hand-linked households, for this sample restriction we define a household member name "match" as having Jaro-Winkler string distance below 0.30 (searching within family members only). In Column 6, we limit the hand-linked records to those with an exact name match and an age difference that is one year or less. All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Appendix Table 4. Effects of SES on Census Racial Assignment, Restricted Samples

	Baseline Handlinks	In Branch or Border County	Exclude NYC Branch	Exclude Foreign-born
	(1)	(2)	(3)	(4)
Panel A: White or Mulatto vs. Black				
Wealth ≥ $100	0.108	0.063	0.117	0.114
	(0.007)	(0.009)	(0.008)	(0.008)
Log Wealth	0.041	0.025	0.051	0.045
	(0.002)	(0.003)	(0.003)	(0.002)
Literate	0.122	0.093	0.120	0.124
	(0.008)	(0.009)	(0.008)	(0.008)
Number of Depositors	14,138	8,827	12,797	13,659
Panel B: White vs. Mulatto or Black				
Wealth ≥ $100	0.068	0.015	0.074	0.072
	(0.005)	(0.004)	(0.005)	(0.005)
Log Wealth	0.030	0.009	0.037	0.033
	(0.002)	(0.002)	(0.002)	(0.002)
Literate	0.067	0.023	0.065	0.067
	(0.005)	(0.004)	(0.005)	(0.005)
Number of Depositors	14,138	8,827	12,797	13,659
Panel C: Mulatto vs. Black				
Wealth ≥ $100	0.085	0.067	0.084	0.085
	(0.009)	(0.010)	(0.009)	(0.009)
Log Wealth	0.043	0.034	0.043	0.043
	(0.004)	(0.004)	(0.004)	(0.004)
Literate	0.092	0.089	0.089	0.092
	(0.008)	(0.010)	(0.008)	(0.008)
Number of Depositors	11,999	7,782	11,649	11,988

Notes: Column 1 reports our baseline estimates from Table 3 (in Columns 2, 4, and 6 of Panels A, B, and C). Column 2 limits the sample to depositors that are in the same county or neighboring county in the 1870 Census as the Freedman Bank branch where they opened their account (between 1866 and 1874). Column 3 excludes all depositors associated with the Freedman Bank branch in New York City, which had the largest share of white depositors (and foreign-born depositors). Column 4 excludes all foreign-born depositors. All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Appendix Table 5. Effects of SES on Census Racial Assignment, Robustness to Alternative Samples

	Baseline Handlinks	No "Racial" Categories	Household Head Only	Only Women	Only Children	Only Wealth ≥ $100
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: White or Mulatto vs. Black						
Wealth ≥ $100	0.108	0.147	0.089	0.123	0.121	-
	(0.007)	(0.011)	(0.010)	(0.016)	(0.013)	
Log Wealth	0.041	0.066	0.036	0.044	0.042	0.041
	(0.002)	(0.004)	(0.003)	(0.005)	(0.004)	(0.004)
Literate	0.122	0.157	0.118	0.135	0.114	0.166
	(0.008)	(0.010)	(0.010)	(0.016)	(0.016)	(0.016)
Number of Depositors	14,138	8,726	7,699	3,952	3,852	3,843
Panel B: White vs. Mulatto or Black						
Wealth ≥ $100	0.068	0.099	0.048	0.090	0.086	0.000
	(0.005)	(0.007)	(0.006)	(0.011)	(0.009)	
Log Wealth	0.030	0.050	0.023	0.037	0.036	0.037
	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)
Literate	0.067	0.084	0.067	0.093	0.060	0.115
	(0.005)	(0.006)	(0.006)	(0.011)	(0.010)	(0.012)
Number of Depositors	14,138	8,726	7,699	3,952	3,852	3,843
Panel C: Mulatto vs. Black						
Wealth ≥ $100	0.085	0.095	0.070	0.089	0.100	-
	(0.009)	(0.011)	(0.011)	(0.019)	(0.018)	
Log Wealth	0.043	0.047	0.036	0.040	0.053	0.049
	(0.004)	(0.005)	(0.005)	(0.007)	(0.008)	(0.008)
Literate	0.092	0.116	0.086	0.095	0.091	0.127
	(0.008)	(0.011)	(0.010)	(0.017)	(0.017)	(0.019)
Number of Depositors	11,999	8,189	6,899	3,387	2,806	2,574

Notes: Column 1 reports our baseline estimates from Table 3 (in Columns 2, 3, and 4 of Panels A, B, and C). Each coefficient is from a separate regression and, as before, the sample size is smaller for Literate. Column 2 omits the more ``racial'' categories of skin tone: (black, nearly black, colored, mulatto, white, nearly white). Column 3 restricts the sample to Household Heads only, Column 4 restricts the sample to women only, and Column 5 restricts the sample to children only (ages 17 and younger). Column 6 restricts the sample to households with measured wealth (≥ $100). All specifications include Freedman Bank branch fixed effects, and robust standard errors are reported in parentheses.

Appendix Table 6. Census Racial Assignment by Literacy of Household Head and Depositor

	White or Mulatto vs. Black		White vs. Mulatto or Black		Mulatto vs. Black	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Household Head Literacy						
Household Head Literacy	0.131	0.131	0.086	0.087	0.096	0.092
	(0.008)	(0.008)	(0.005)	(0.005)	(0.008)	(0.008)
Skin Tone Index	-0.279		-0.269		-0.303	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999
Panel B: Depositor Literacy						
Depositor Literacy	0.124	0.122	0.065	0.067	0.094	0.092
	(0.008)	(0.008)	(0.005)	(0.005)	(0.008)	(0.008)
Skin Tone Index	-0.283		-0.271		-0.301	
	(0.004)		(0.003)		(0.011)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	13,029	13,029	13,029	13,029	11,277	11,277
Panel C: Household Head and Depositor Literacy						
Household Head Literacy	0.085	0.087	0.079	0.077	0.055	0.050
	(0.013)	(0.013)	(0.007)	(0.007)	(0.014)	(0.014)
Depositor Literacy	0.062	0.058	0.007	0.010	0.055	0.056
	(0.012)	(0.012)	(0.006)	(0.006)	(0.013)	(0.013)
Skin Tone Index	-0.279		-0.267		-0.300	
	(0.004)		(0.003)		(0.011)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	13,029	13,029	13,029	13,029	11,277	11,277

Notes: This Table reports estimates that correspond to depositor literacy in Table 3, but for the literacy of the Household Head (in Panel A) and for the jointly estimated effects of Household Head literacy and depositor literacy (in Panel C).

Appendix Table 7. Census Racial Assignment by Measured Personal Property and Real Estate

	White or Mulatto vs. Black		White vs. Mulatto or Black		Mulatto vs. Black	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Measured Real Estate Property						
Real Estate Value	0.140	0.136	0.089	0.094	0.121	0.117
	(0.010)	(0.009)	(0.007)	(0.007)	(0.013)	(0.012)
Skin Tone Index	-0.289		-0.276		-0.306	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999
Panel A: Measured Personal Property						
Personal Property	0.103	0.105	0.084	0.083	0.077	0.074
	(0.008)	(0.008)	(0.006)	(0.006)	(0.010)	(0.010)
Skin Tone Index	-0.290		-0.274		-0.309	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999
Panel C: Measured Real Estate and Personal Property						
Real Estate Value	0.113	0.107	0.063	0.068	0.103	0.100
	(0.010)	(0.010)	(0.007)	(0.007)	(0.013)	(0.013)
Personal Property	0.068	0.072	0.065	0.062	0.052	0.050
	(0.009)	(0.008)	(0.005)	(0.005)	(0.010)	(0.010)
Skin Tone Index	-0.284		-0.271		-0.304	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999

Notes: This Table reports estimates that correspond to measured wealth in Table 3, but separating total household wealth into real estate wealth (in Panel A), personal property wealth (in Panel B), and the jointly estimated effects (in Panel C).

Appendix Table 8. Census Racial Assignment by Log Personal Property and Real Estate

	White or Mulatto vs. Black		White vs. Mulatto or Black		Mulatto vs. Black	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Log Real Estate Wealth						
Log Real Estate Wealth	0.039	0.039	0.029	0.029	0.048	0.047
	(0.003)	(0.003)	(0.002)	(0.002)	(0.005)	(0.005)
Skin Tone Index	-0.287		-0.272		-0.303	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999
Panel A: Log Personal Property						
Log Personal Wealth	0.043	0.046	0.046	0.043	0.045	0.044
	(0.003)	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)
Skin Tone Index	-0.287		-0.267		-0.308	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999
Panel C: Log Real Estate and Personal Property						
Log Real Estate Wealth	0.029	0.028	0.015	0.016	0.040	0.039
	(0.003)	(0.003)	(0.002)	(0.002)	(0.005)	(0.005)
Log Personal Wealth	0.024	0.029	0.036	0.033	0.027	0.027
	(0.004)	(0.004)	(0.003)	(0.003)	(0.006)	(0.006)
Skin Tone Index	-0.282		-0.265		-0.302	
	(0.003)		(0.003)		(0.010)	
Skin Tone Fixed Effects		Yes		Yes		Yes
Number of Depositors	14,138	14,138	14,138	14,138	11,999	11,999

Notes: This Table reports estimates that correspond to log wealth in Table 3, but separating total household wealth into real estate wealth (in Panel A), personal property wealth (in Panel B), and the jointly estimated effects (in Panel C). For households with less than $100 in measured wealth, we continue to assign a value of $45.25 before taking the natural log (based on the average among the 2% of households nationwide for which Census enumerators did report wealth between $0 and $100).

Appendix Table 9. Effects of SES on Census Racial Assignment, Alternative Specifications

			Additional controls for:			
	Baseline Handlinks	No Branch FE	Branch by Acc. Open Year FE	Emancipated by 1860	Age and Gender	County FE
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: White or Mulatto vs. Black						
Wealth ≥ $100	0.108	0.102	0.108	0.105	0.111	0.102
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Log Wealth	0.041	0.040	0.041	0.040	0.042	0.037
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Literate	0.122	0.124	0.121	0.119	0.123	0.118
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Number of Depositors	14,138	14,138	14,138	14,138	14,138	14,138
Panel B: White vs. Mulatto or Black						
Wealth ≥ $100	0.068	0.066	0.069	0.071	0.069	0.045
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Log Wealth	0.030	0.030	0.030	0.031	0.031	0.022
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Literate	0.067	0.069	0.067	0.069	0.068	0.055
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Number of Depositors	14,138	14,138	14,138	14,138	14,138	14,138
Panel C: Mulatto vs. Black						
Wealth ≥ $100	0.085	0.078	0.084	0.079	0.088	0.091
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Log Wealth	0.043	0.041	0.043	0.040	0.044	0.043
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Literate	0.092	0.092	0.091	0.087	0.092	0.097
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Number of Depositors	11,999	11,999	11,999	11,999	11,999	11,999

Notes: Column 1 reports our estimates from Table 3 (in Columns 2, 4, and 6 of Panels A, B, and C). Columns 2 to 6 report estimates for alternative specifications with different controls. Column 2 omits the branch fixed effects. Column 3 includes fixed effects for each branch interacted with the bank account opening year. Columns 4 to 6 are the same as our baseline specification, except for the addition of: an indicator variable for whether the depositor was free in 1860 (restricted to those racialized as Black or Mulatto in the 1870 Census); an indicator for gender and controls for age and age-squared; and fixed effects for the depositor's county in the 1870 Census. All specifications include Freedman Bank branch fixed effects (except Column 2), and robust standard errors are reported in parentheses.

Appendix Table 10. Differences in Bank Record Characteristics by Skin Tone, Attenuated Within Census Race

			Within Census Race:		
	Full Sample		Black	Mulatto	White
	(1)	(2)	(3)	(4)	(5)
Panel A: Blackness of Name					
Skintone Index	0.158	0.086	0.177	0.098	0.072
	(0.007)	(0.012)	(0.026)	(0.027)	(0.012)
Black		0.243			
		(0.031)			
Mulatto		0.093			
		(0.031)			
Number of Depositors	14,057	14,057	9,899	2,027	2,131
Panel B: Professional Occupation in Depositor Records					
Skintone Index	-0.085	-0.068	-0.053	-0.057	-0.075
	(0.005)	(0.008)	(0.011)	(0.018)	(0.011)
Black		-0.056			
		(0.020)			
Mulatto		-0.030			
		(0.020)			
Number of Depositors	10,446	10,446	7,880	1,387	1,179
Panel C: Writes Own Signature in Depositor Book					
Skintone Index	-0.158	-0.141	-0.137	-0.144	-0.208
	(0.004)	(0.008)	(0.015)	(0.019)	(0.010)
Black		-0.055			
		(0.023)			
Mulatto		0.039			
		(0.022)			
Number of Depositors	12,110	12,110	8,579	1,755	1,776

Notes: This Table reports estimates that correspond to those in Table 7, but for the depositor Bank record characteristics that we report in Table 6.