

Census Race Categories and the Mortality of Mixed-Race American Men in the 20th Century*

Maria Osborne[†]

February 27th, 2023

Abstract

Despite substantial research on racial disparities in mortality, little is known about the mortality experiences of mixed-race and multiracial groups in the United States. In this paper, I use census data from the early 1900s, which included mulatto as a race category, linked to Social Security Administration death records to study the all-cause mortality of black, white, and mixed black/white men. I find that in the aggregate, all men of African descent, both mixed-race and single race, experience worse mortality than whites. Additionally, mulattos may have a mortality advantage over blacks. Within particular social groups, however, the mortality outcomes of mulattos often are not meaningfully distinct from other racial groups.

* The author would like to acknowledge financial support from NIA grant R01AG058940. Additional thanks to Casey F. Breen, Sharon Green, Serge Atherwood, Andrea Miranda Gonzalez, Joshua R. Goldstein, and Nathan Seltzer for their valuable comments and suggestions.

[†] Department of Demography, University of California, Berkeley. mariaosborne@berkeley.edu

I. Introduction

In North American contexts, research on race and mortality rests on the assumption that racial identities comprise stable, mutually exclusive categories. This is not without justification: for much of the last century, the dominant American philosophy of race has been one of static and singular identity (Nobles, 2000). Data collection which neatly sorts people into distinct racial classifications both reflects and perpetuates this ideology. Yet, recent changes to the U.S. decennial census indicate a shift towards the definition and recognition of mixed race and multiracial individuals. For the first time, in the year 2000, respondents could select more than one race – and over 6 million Americans did (Jones & Symens Smith, 2001). But this is hardly the first time the “official” boundaries of racial categorization have changed, and not even the first time that the census has attempted to measure populations of mixed racial heritage. Indeed, racial categories specifically for persons of mixed black and white ancestry existed on census forms from 1850-1920. In most years this category was called “mulatto”; the 1900 census instead used the wording “of Negro descent,” and the 1890 census included “octoroon” and “quadroon” alongside “mulatto” as a further attempt to quantify black blood quantum (Humes & Hogan, 2009). Persons of mixed African and European heritage existed in the Americas long before the U.S. government decided to explicitly count them. The inclusion of mulatto as a specific racial designation was not so much an indicator of demographic changes as it was the result of race science, and the political influence of race science (Nobles, 2004). Its eventual removal was also a politically motivated process borne from the rise of the one-drop philosophy, which strictly delineated whites from any others with non-white “blood,” as well as perceived unreliability of mulatto enumeration (Saperstein & Gullickson, 2013).

These paradigms of firm racial boundaries and hypodescent (classification of individuals with mixed ancestry as a member of their lowest socially ranked racial group) became enshrined in both American law and popular imagination. Even within this system of discrete racial categories, however, changes both in racial self-identification and observed racial classification can occur (Saperstein & Penner, 2012). The connections between mortality and subtler dimensions of race, such as multiracial

heritage and diversity of skin color, are rarely studied. Green and Hamilton (2013) studied mortality outcomes among mulattos in late 19th century post-reconstruction North Carolina, and found that both blacks and mulattos experienced higher mortality than whites. Results on mulattos compared to blacks depended on gender; only mulatto women had a significant mortality advantage over other black women. Studies using more recent data find evidence of a skin color gradient in mortality but tend to be geographically limited in scope (Knapp et al., 1995; Keil et al., 1992). In perhaps the most nationally representative study of color and mortality among blacks, Stewart et al. (2020), find that lighter skin correlates with lower mortality hazards, but that returns to skin color vary upon dimensions like educational attainment. Forms of morbidity among multiracial persons do not always imply a simple relationship between race or skin color and health. Veenstra (2019), for example, finds that in a contemporary Canadian population, mixed black/white persons have worse mental health than blacks, but lower rates of hypertension.

This paper investigates all-cause mortality of mulattos born in the early 20th century. I link cohorts of male children born from 1900-1920 who were identified as mulatto on their first census (either 1910 or 1920) to the 1940 census (where the mulatto category did not exist) and then to Social Security Administration death records. This allows for two comparisons: first, the mortality of persons identified as mulatto as children relative to mortality of those identified as black or white. Secondly, I look for differences between mulatto children who as adults were classified as black versus those who were classified as white. This latter point of comparison sheds light on the ways that racial classification changes over time, heterogeneity of life and mortality experiences among multiracial groups, and the way conclusions about race are informed by the way it is measured. The conclusions drawn here are also twofold. For one, I am interested in the mortality outcomes themselves of single race and mixed-race African Americans. These outcomes can then be informative about the meaning and function of race, and specifically of blackness, in 20th century American society. A key aspect of this analysis is that it does not use a measure of racial self-identification, but a racial classification determined by census enumerators. This study, therefore, explores the links between mortality and the social assignment of race.

1a. Background: Theories of race, racism, and mortality

Racial disparities in mortality in the United States are a well-known phenomenon. Across a wide variety of metrics, African Americans have persistently been shown to have worse mortality outcomes than white Americans. A substantial gap in life expectancy at birth between whites and non-whites has been documented throughout the 20th century: in 1900 this gap was a staggering 14 years for men and 13 years for women (Arias & Xu, 2019). Prior to the COVID-19 pandemic, the life expectancy gap between whites and blacks narrowed to 3.6 years by 2018 but has been relatively stagnant for several years (Schwandt et al., 2021). There are several specific contributors to black/white mortality differences, including higher infant mortality related to short gestation and low birth weight (Ely & Driscoll, 2020), greater infectious disease burden (Richardus, 2001), and higher mortality burden from some chronic diseases, including hypertension and diabetes (Wong et al., 2002).

The causal mechanisms driving disproportionately poor health outcomes among blacks are not yet completely understood. General racial differences in socioeconomic status are certainly related: black Americans fall behind whites in numerous measures that are predictive of health and mortality, such as income and education (Geruso, 2012). Geruso estimates that standard socioeconomic and demographic indicators including income, education, and homeownership explain around 80% of the contemporary (pre-COVID) gap in life expectancy at age one for men, and 70% of the gap for women. Geronimus (1992) proposed the *weathering hypothesis* as an explanation for the racial disparity in infant mortality, which posits that black women's health deteriorates earlier than that of non-Hispanic white women as a "physical consequence of cumulative socioeconomic disadvantage". Further research on this hypothesis has emphasized the deleterious physical consequences of experiencing the stress of racial stigmatization (Geronimus, 2006). Hummer (1996) argues for a model of racial mortality disparities that considers the role of racism at both the individual and systemic level. Discounting genetic or biological definitions of race, Hummer emphasizes the historical and social role of race as the root cause of socioeconomic disadvantages and behavioral differences (e.g., smoking), from which health disparities may arise.

Particularly in the study of mortality, a binary delineation between black and white is often taken at face value. This is not necessarily correct nor incorrect; however, a look into the history of mulattos reveals alternate paradigms of race that have existed in the Americas. People with mixed black and white parentage have existed since the transatlantic slave trade. While the term “mulatto” originally referred to a child of one “purely” black and one “purely” white parent, by at least the mid 19th century it was usually meant to encompass all people with identifiable white and black ancestry. Partially because some mulattos were born free and others enslaved, mulattos occupied a range of social positions in the antebellum south. Some were favored slaves (often working within households or as skilled craftsmen rather than as agricultural laborers), some were considered free blacks, and others were considered legally white. While their status varied somewhat according to local customs and laws, in certain places the social standing of mulattos may have existed somewhere between that of enslaved blacks and free whites (Davis, 1991). Various metrics can quantify possible advantages of mulattos over blacks. For example, mulattos had greater adult height (an indicator of childhood health) than other blacks in the antebellum South (Bodenhorn, 2002), and greater household wealth (Bodenhorn, 2006) in the reconstruction era. In at least some areas of the post-slavery South, a level of class and occupational distinction separated blacks, mulattos, and whites into a three-tiered hierarchy (Gullickson, 2010). There are two avenues through which it is theorized that mulattos gained social and economic privileges: lighter skin color, and privileges (e.g., wealth) inherited from white/free family members. Discrimination based on darkness of skin tone is arguably the more important factor (Hill, 2000). While difficult to generalize, this history of class and race suggests possible advantages in morbidity and mortality among mulattos.

Any popular notion of mulattos as meaningfully racially distinct from others of black ancestry had largely broken down by the early 20th century—an era where the eugenics movement gained popularity, one-drop rules became increasingly codified, and Jim Crow laws encouraged the definition of rigid color lines in service of enforcing racial hierarchies (Gullickson, 2010). This does not mean that black and multiracial America abandoned ideas of racial mixture and ambiguity. Writer Langston Hughes, who was of white, black, and Cherokee lineage, wrote: "You see, unfortunately, I am not black.

There are lots of different kinds of blood in our family. But here in the United States, the word ‘Negro’ is used to mean anyone who has any Negro blood at all in his veins” (Hughes, 1940). Hughes’ contemporary Nella Larsen, herself of mixed Danish and Afro-Caribbean parentage, explored in her 1929 novel *Passing* the lives of two white-passing black women in 1920’s New York City. The book highlights some aspects of the lived experience of light-skinned blacks: namely, the ability to ingratiate oneself within mainstream white society, and the anxieties stemming from a lack of belonging in any one racial group.

Both these examples serve to illustrate the multifaceted nature of race, an idea expounded upon in modern sociological scholarship. Roth (2016) notes that the concept of race does not constitute a single identity, but is a multidimensional phenomenon that includes phenotype, observed race, self-classification, ancestry, and more. Observed race is the basis for racial discrimination, and in this sense data from pre-1960 censuses comprise a uniquely useful measure of race in the context of mortality since racial classification was assigned rather than self-selected. If individual-level racism plays a part in shaping health outcomes, then the way a person is assigned a particular race by an outsider should give some insight into their social position and experience of discrimination. Relying on this measure of observed race connects this study to a growing body of literature that explores the ways in which observed skin color matters, even within of mutually exclusive race groupings. Among blacks, skin color is a predictor of occupation and income (Keith & Herring, 1991). Darker skinned people experience colorism: greater discrimination and worse treatment than their lighter peers due their appearance (Monk, 2014).

The dynamic and unique history of mulattos in relation to privilege and power, alongside contemporary research on colorism, offers unique insights into the socially constructed nature of race. It also calls into the question the tendency for mortality research to treat racial categories as both prescriptive and homogeneous. In light of these histories and theories of race, two possible hypotheses about the relationship between mixed-race status and mortality emerge. First, mortality may exist on a continuum, whereby mixed-race and lighter-skinned African Americans experience mortality on a level

somewhere between that of single-race blacks and whites. In this scenario mixed-race persons, facing less discrimination and more material/financial capital, “accumulate” less discrimination over their life course and have longer lives. However, since they are still marked as black in some degree, they do not have the same mortality advantage as whites. An alternative hypothesis is that the dogma of the one-drop rule, enforced by white mainstream political and social structures, translates to mortality and that all people marked as black, regardless of racial ancestry or skin color, experience similar mortality risks.

2. Data and Methods

2a. Data sources

I use IPUMS-USA microdata from the decennial censuses of 1910, 1920, and 1940 (Ruggles et al., 2021), linked using crosswalks from the Census Linking Project (Abramitzky et al., 2020). The sample is limited to U.S.-born males classified as white, black, or mulatto on census records. Mortality data come from the CenSoc-DMF file (Goldstein et al., 2021), which matches 1940 census records to the Social Security Death Master File. Death records in the CenSoc-DMF are available from 1975-2005. I only include deaths at age 65 or later. Although the mulatto race category was not included on the 1940 census, linking these earlier census records to 1940 is useful in several ways. Foremost, CenSoc datasets that connect census and mortality records are specifically constructed using the 1940 census. Secondly, it allows us to observe what becomes of mulatto children as adults, including measures such as their educational attainment. Thirdly, it provides an additional observation of race under a different racial classification schema: former mulattos could not be categorized as such in 1940, and their adult racial classification could help illuminate relationships between color and mortality. Combining these sources produces a dataset that is large, nationally representative, and rich in individual-level data. All links between data sources utilize a version of the ABE matching algorithm that standardizes and cleans name spellings, allows for some flexibility on birth year, and requires matches to be unique within a four-year window. This latter requirement is meant to ameliorate the issue of false matches. The data set contains N=930,334 records. 422,756 men are initially identified in the 1910 census and 507,578 are initially identified in the

1920 census. Table 1 provides an overview of the socioeconomic and demographic characteristics of the sample, divided into groups based on racial classification on a person's first census.

2b. Census race variables

Decennial censuses before 1960 were carried out by trained door-to-door enumerators, who were tasked with determining the racial status of Americans. Instructions to enumerators for the 1910 census include directions to mark “evidently full-blooded Negroes” as black, whereas mulattos are persons with “some proportion or perceptible trace of Negro blood” (US Department of Commerce, 1910). Instructions for the 1920 census are similar, explaining that: “the term ‘black’ (B) includes all Negroes of full blood, while the term ‘mulatto’ (Mu) includes all Negroes having some proportion of white blood” (US Department of Commerce, 1919). Ostensibly, then, the distinction between black/Negro and mulatto was based on blood quantum. In practice, little is known about how enumerators differentiated blacks and mulattos. Their classification may have been based on perceived appearance, family relationships, community status, occupation, local customs, and the personal convictions of enumerators. The mulatto/black distinction presumably reflects the skin color of individuals and household members more than other factors (Gullickson, 2010). The wording of the 1910 census instructions to enumerators, with its use of terminology such “evidently” and “perceptible trace of negro blood”, implies that perceived appearance was indeed an acceptable metric for determining race. I operate under the assumption that people designated as mulatto were, on average, lighter skinned than those designated as black.

The 1940 census did away with this ambiguity, telling enumerators that “A person of mixed white and Negro blood should be returned as a Negro, no matter how small the percentage of Negro blood. Both black and mulatto persons are to be returned as Negroes, without distinction” (U.S. Department of Commerce, 1940). These instructions reveal that while the word “mulatto” may have been a recognizable and even potentially meaningful term in 1940, that it was no longer considered by the federal government to constitute a distinct group.

2c. Analytical strategy

I use multiple linear regression to predict age at death from race and other variables. Since only deaths at age 65 or above are included, the dependent variable (age at death in years) is conditional on survival to age 65. Additionally, mortality data are doubly truncated, and therefore the death ages observed for each cohort are different. All models include birth year fixed effects to account for the effect of changing observational windows and are of the form:

$$\text{Age at death} = \beta_0 + \beta\mathbf{X} + \gamma_t t + \epsilon$$

where \mathbf{X} is a vector of covariates, including a dummy variable for race, and $\gamma_t t$ are birth year fixed effects. I use logistic models to explore socioeconomic and contextual dimensions of shifts between census racial classifications.

3. Results

3a. Mortality based on a single measure of race

Table 2 compares the longevity of men based racial classification on their first census, either 1910 or 1920 depending on their birth year. Model 1 includes only race and birth year as explanatory variables, as birth year must be included to account for the structure of the data. Men designated as either mulatto or black as children had decreased later-life longevity compared to men designated as white. Relative to whites, the reduction in lifespan for blacks (-0.838 years, 95% CI = (-0.941, -0.736)) is greater in magnitude than the reduction for mulattos (-0.558 years, 95% CI = (-0.730, -0.386)). Model 2 introduces additional variables, including standard SES indicators such as educational achievement (measured here as a dichotomous indicator of high school completion) and income as measured in the 1940 census. Remaining life expectancy for blacks remains lower than the white reference group, but the results for mulattos are more ambiguous, not clearly differentiated from either blacks or whites.

High school completion is strongly associated with increased longevity, and Table 3 dives deeper into this dimension, comparing the longevity of racial groups at the same level of educational attainment

(e.g, blacks and mulattos who have completed high school are compared to a reference group of whites who have completed high school). At any given level of education, there is not significance evidence that mulattos as a group have distinct mortality experiences from both whites and blacks. Mulattos tend to have reduced longevity compared to whites, and the effect size is comparable to that of blacks, except in model 4 (non-high school graduates), which again is more ambiguous. Additionally, there are larger white/non-white disparities among high school graduates than among those who did not graduate high school.

3b. 1940 racial classification and mortality

Mulattos are necessarily sorted into some other racial group on the 1940 census, when that category no longer existed. The majority of mulatto children (about 85%) were identified as black in the 1940 census, and the rest as white. (While some mulattos had a different racial classification in 1940, my sample excludes all race categories other than black, white, and mulatto.) For the analysis in Table 4, I break the sample into discrete categories based on the combination of their racial designation on their first census and the 1940 census and use these categories as dummy variables in regression on age at death. Mulattos are included in two of these categories: those who later were classified as black (denoted “mulatto-to-black”), and those who changed from mulatto to white (denoted “mulatto-to-white”). The mortality experiences of these two groups are shown to be distinct. Compared to the reference group of men who were categorized as white on both their first childhood census and the 1940 census, the mulatto-to-black group had reduced longevity (0.673 fewer years of life after 65, 95% CI = (-0.859, -0.486)), comparable to that of men consistently classified as black. Mulatto-to-white men did not see a reduction in longevity compared to the reference group.

The racial classification of (former) mulattos in 1940 is related to a number of sociodemographic variables. Log odds ratios of factors predicting race category of on the 1940 census are presented in Figure 1. Being a high school graduate, higher occupational income score in 1940, being currently married in 1940, and older age all increase the likelihood that someone formerly classed as mulatto will

be classed as white on the 1940 census. Conversely, urban residence in 1940, being born or residing in the Deep South, and cohabitating with a parent in 1940 increase odds that a former mulatto will be categorized as black on the 1940 census. Some of these models point to higher SES among the mulatto-to-white group.

An important caveat to these results is that connecting multiple data sets inevitably introduces false matches into the data. The effect seen in Table 4, where the mulatto-to-white group experiences comparable mortality as whites, could be the result of mulattos being incorrectly linked to whites in the 1940 census. One way to assess this is to compute rates of middle initial agreement between census matches. Middle initial was requested on both the 1910 and 1920 census population schedules, though not in the 1940 census. Even so, a portion of 1940 census records do contain some information about middle name. About 15% of the dataset contains middle initial or name in both the initial census record and the 1940 census record. Table 5 shows the middle initial agreement rates between censuses. The agreement rate for white-to-white, black-to-black, and mulatto-to-black are all above 86%, while the mulatto-to-white agreement rate is much lower, 66%. This suggests that the mulatto-to-white group is more likely to contain false matches than the other groups, which could distort results by making this group look more like whites than it is.

4. Discussion

Using data from early 20th century censuses, I find that male children designated as mulatto or black on their first census experienced decreased longevity compared to white children, conditional on survival to age 65. Additionally, considering no other dimensions besides the child's first racial designation, mulattos are at a mortality advantage compared to blacks. Results from Table 2 therefore support the idea of mulattos forming a sort of intermediate class between blacks and whites. However, this evidence for this difference not overwhelmingly strong, as the 95% confidence intervals for the effect sizes are very nearly overlapping. These results echo the findings of Green and Hamilton (2013), which found a mortality disadvantage (in terms of the mortality rate at age 30) for all non-white males compared

to whites in post-reconstruction North Carolina, but weaker evidence for a mortality difference between mulattos and blacks.

Controlling for basic socioeconomic factors suggests possible explanations for this modest distinction between the two race groups. At a given level of education, there is no statistically significant difference between blacks and mulattos. Indeed, in models that stratify by education and/or control for other SES factors, the longevity level of mulattos no longer falls cleanly between that of whites and blacks. Again, this result has precedence in the literature: Stewart et al. (2020) found that at the same level of education, medium-skinned and dark-skinned blacks did not have differential levels of mortality from each other. While the measures of observed skin color in that study (light, medium, and dark) do not precisely map onto the measure of race the I use, the finding is similar: within important social categories like educational groups, gradations in race become less distinct. However, race may influence the ability to enter these sorts of groups. It is important to note that controlling for socioeconomic factors should not be done without acknowledging the role of systemic racism in giving rise to differing levels of socioeconomic resources by race (Hummer, 1996). As shown in Table 1, Both mulattos and blacks are far behind whites in rates of educational attainment, though mulattos have higher rates of high school completion than blacks. The small mortality advantage of mulattos could be connected to their ability to incur more education: either through the effects of education itself, or related social advantages.

Incorporating racial classification in 1940, which may hint at phenotypical differences within the mulatto group itself, shows that only the group of mulattos later categorized as white had better mortality outcomes than people consistently categorized as black. The majority, who were designated to be black as adults, had longevity not significantly different than those always identified as black. This suggests that most African American men, regardless of their skin color or racial heritage, experienced comparable levels of mortality. It is unsurprising that most mulattos later came to be identified as black, as the mulatto category was dropped from the census and the term “black” came to broadly encompass all peoples of African descent. Mortality appears to follow this blunt distinction. This mulatto-to-white group are a more unusual group who census enumerators would have recognized as persons of color in

childhood, but *not* as adults. While we cannot measure skin color directly, having relatively light skin is surely related to this phenomenon. Context, however, is also key: mulattos born or residing in the south are less wont to “become” white in the eyes of census enumerators, for example. So are those who live with at least one parent, perhaps because the presence of a more obviously non-white parent would in turn mark that person as non-white. Education and occupation are both linked to whitening. Status, family/community, and geography are clearly connected to certain shifts in racial classification.

These results on race shifting between censuses are quite tenuous, as they rely on the assumption that individuals are being correctly linked between data sets. If they are, then this is evidence that mulattos form two distinct groups based on their ability to be classified as white by observers. The “in between” nature of mortality for mulattos in aggregate could then be driven by a mortality advantage for a select privileged few. Alternatively, if the mulatto-to-white group is made up of a high proportion of mismatched individuals, making them appear better off than they are in reality, then the mortality advantage of mulattos over blacks may be driven by data errors. In both cases, a similar conclusion arises: most of the group designated as mulatto has longevity levels comparable to that of blacks. Broadly, if mulattos are disaggregated in any way, for instance by SES level or later racial categorization, they look less like a meaningfully distinct group in terms of mortality.

There are a number of limitations to this study. The conclusions drawn here are not causal, and thus the specific pathways between race and mortality remain somewhat obscure. There are also potential issues surrounding the measurement of race on pre-1960 censuses. Although the distinctions between black/mulatto and mulatto/white are likely related to skin color, it is impossible to know the standards and idiosyncrasies of individual census enumerators. For instance, according to Table 1, mulattos were more likely to be born in the northern and western regions than blacks, while blacks were more concentrated in the South than mulattos. It is possible that difference arises because southern census enumerators understood and perceived race differently than northern enumerators, rather than any substantive differences in ancestry or appearance among northern and southern African Americans. It is also possible that census enumerators of different races observed race differently. The Census Bureau

believed that an increase in mulattos on the 1910 census compared to previous censuses was due to higher employment of black census enumerators (Nobles, 2000). This implies that white enumerators were more likely to see race in strictly black or white terms—literally. Had the censuses been carried out strictly by whites, it is possible that more distinct differences between the black and mulatto category would emerge, if whites only perceived very select, relatively privileged, African Americans as mulatto rather than black.

This paper investigates racial differences in mortality by considering facets of race such as ancestry and phenotype that are discussed in sociology but often ignored in mortality literature. While the mulatto category is problematic to work with on several levels, it highlights the multidimensional aspects of race and the politics and limitations of racial categorization schema. As self-identification with multiple races becomes more common, as occurred between the 2000 and 2010 census, mortality research needs to contend with changing paradigms of race and the assumptions embedded in existing racial classification systems.

5. References

- Abramitzky, R., Boustan, L., & Rashid., M. (2020). *Census linking project* (Version 1.0) [dataset]. <https://censuslinkingproject.org>
- Arias, E., & Xu, J. (2019). United States life tables, 2017. *National vital statistics reports*, 68(7). 1-65.
- Bodenhorn, H. (2002). The Mulatto advantage: The biological consequences of complexion in rural antebellum Virginia. *Journal of interdisciplinary history*, 33(1). 21–46.
- Bodenhorn, H. (2006). Colorism, complexion homogamy, and household wealth: Some historical evidence. *The American economic review*, 96(2). 256–260.
- Davis, J. F. (1991). *Who is black: One nation's definition*. Pennsylvania State University Press.
- Ely, D.M., & Driscoll, A.K. (2020). Infant mortality in the United States, 2018: Data from the period linked birth/infant death file. *National vital statistics report*, 69(7). National Center for Health Statistics.
- Geronimus, A.T. (1992). The weathering hypothesis and the health of African-American women and infants: evidence and speculations. *Ethnicity & disease*, 2(3). 207–221.
- Geronimus, A.T., Hicken, M., Keene, D., & Bound, J. (2006). “Weathering” and age patterns of allostatic load scores among blacks and whites in the United States. *American journal of public health*, 96(5). 826-833.
- Geruso, M. (2012). Black-white disparities in life expectancy: how much can the standard SES variables explain?. *Demography*, 49(2), 553–574.
- Goldstein, J. R., Alexander, M., Breen, C., Miranda González, A., Menares, F., Osborne, M., Snyder, M., & Yildirim, U. (2021) *CenSoc Mortality File* (Version 2.0) [dataset]. Berkeley: University of California.
- Green, T.L., & Hamilton, T.G. (2013). Beyond black and white: color and mortality in post reconstruction era North Carolina. *Explorations in economic history*, 50(1), 148-159.
- Gullickson, A. (2010). Racial boundary formation at the dawn of Jim Crow: the determinants and effects of black/mulatto occupational differences in the United States, 1880. *American journal of sociology*, 116(1), 187–231.
- Hill, M. E. (2000). Color differences in the socioeconomic status of African American men: Results of a longitudinal study. *Social forces*, 78(4), 1437-1460.
- Hughes, L. (1940). *The big sea: An autobiography by Langston Hughes*. Alfred A. Knopf.
- Humes, K., & Hogan, H. (2009). Measurement of race and ethnicity in a changing, multicultural America. *Race and social problems*, 1(3). 111-131.
- Hummer, R.A. (1996). Black-white differences in health and mortality: A review and conceptual model. *The sociological quarterly*, 37(1). 105-125.
- Jones, N. A., & Symens Smith, A. (2001). *The two or more races population: 2000*. United States Census Bureau.
- Keil, J. E., Sutherland, S. E., Knapp, R. G., Tyroler, H. A., & Pollitzer, W. S. (1992). Skin color and mortality. *American journal of epidemiology*, 136(11). 1295–1302.
- Keith, V. M., & Herring, C. (1991). Skin tone and stratification in the black community. *American journal of sociology*, 97(3). 760-778.
- Knapp, R. G., Keil, J. E., Sutherland, S. E., Rust, P. R., Hames, C., & Tyroler., H. A. (1995). Skin color and cancer mortality among black men in the Charleston Heart Study.” *Clinical Genetics*, 47. 200–206.
- Larsen, N. (1929). *Passing*. Knopf.

- Mason, P.L. (2013). *Encyclopedia of race and racism* (2nd ed.). Macmillan Reference USA.
- Monk E. P., Jr (2015). The cost of color: Skin color, discrimination, and health among African-Americans. *American journal of sociology*, 121(2), 396–444.
- Nobles, M. (2000). *Shades of citizenship: Race and the census in modern politics*. Stanford University Press.
- Nobles, M. (2004). Racial categorization and censuses. In D. I. Kertzer & D. Arel (Eds.), *Census and identity: The politics of race, ethnicity, and language in national censuses* (pp. 43- 70). Cambridge University Press.
- Richardus, J.H., & Kunst, A.E. (2001). Black-white differences in infectious disease mortality in the United States. *American Journal of Public Health*, 91(8), 1251-1253.
- Rockquemore, K., & Brunsma, D.L. (2002). Socially embedded identities: theories, typologies, and processes of racial identity among black/white biracials. *Sociological Quarterly*, 43(3), 335-356.
- Roth, W.D. (2016) The multiple dimensions of race. *Ethnic and racial studies* 39(8), 1310-1338.
- Ruggles, S., Flood, S., Foster, S., Goeken, R., Pacas, J., Schouweiler, M., & Sobek, M. (2021). IPUMS USA (Version 11.0) [dataset]. Minneapolis: University of Minnesota
- Saperstein, A., & Gullickson, A. A. (2013). “Mulatto escape hatch” in the United States? Examining evidence of racial and social mobility during the Jim Crow era. *Demography*, 50, 1921–1942.
- Saperstein, A., & Penner, A. M. (2012). Racial fluidity and inequality in the United States. *The American journal of sociology*, 118(3), 676-727.
- Stewart, Q. T., Cobb, R. J., & Keith, V. M. (2020). The color of death: race, observed skin tone, and all-cause mortality in the United States. *Ethnicity & health*, 25(7), 1018–1040.
- U.S. Department of Commerce, Bureau of the Census. (1940). *Instructions to Enumerators*. Government Printing Office.
- U.S. Department of Commerce and Labor, Bureau of the Census. (1910). *Instructions to Enumerators*. Government Printing Office.
- U.S. Department of Commerce, Bureau of the Census. (1919). *Instructions to Enumerators*. Government Printing Office.
- Veenstra, G. (2019). Black, white, black and white: Mixed race and health in Canada. *Ethnicity & health*, 24(2), 113-124.
- Wong, M. D., Shapiro, M. F., Boscardin, W. J., & Ettner, S. L. (2002). Contribution of major diseases to disparities in mortality. *The New England journal of medicine*, 347(20), 1585–1592.

Table 1*Socioeconomic and demographic characteristics of the unweighted sample*

Race on First Census	White N = 904,317 ¹	Black N = 19,306 ¹	Mulatto N = 6,711 ¹
Race in 1940			
White	902,641 (100%)	1,786 (9.3%)	995 (15%)
Black	1,676 (0.2%)	17,520 (91%)	5,716 (85%)
Region of Birth ²			
Midwest	412,882 (46%)	951 (4.9%)	594 (8.9%)
Northeast	218,581 (24%)	898 (4.7%)	452 (6.7%)
South	200,326 (22%)	17,365 (90%)	5,593 (83%)
West	72,514 (8.0%)	92 (0.5%)	72 (1.1%)
Region in 1940 ²			
Midwest	385,930 (43%)	2,417 (13%)	1,150 (17%)
Northeast	212,307 (23%)	2,044 (11%)	866 (13%)
South	188,024 (21%)	14,430 (75%)	4,441 (66%)
West	118,056 (13%)	415 (2.1%)	254 (3.8%)
Urbanicity in 1940			
Rural	425,620 (47%)	10,157 (53%)	3,077 (46%)
Urban	478,697 (53%)	9,149 (47%)	3,634 (54%)
Finished High School	364,062 (41%)	2,449 (13%)	1,318 (20%)
Occupational Income Score ³	23 (14, 27)	18 (11, 20)	20 (14, 23)
Marital Status 1940			
Married	582,267 (64%)	13,422 (70%)	4,703 (70%)
Separated/Divorced/Widowed	10,537 (1.2%)	378 (2.0%)	153 (2.3%)
Single	311,513 (34%)	5,506 (29%)	1,855 (28%)

¹n (%); Median (IQR)

²IPUMS regions consist of the following states: Northeast (Connecticut, Maine, Massachusetts, New Jersey, New Hampshire, New York, Pennsylvania, Rhode Island, Vermont); Midwest (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin); South (Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia); and West (Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, Oregon, New Mexico, Utah, Washington, Wyoming).

³Occupational Income Score is a constructed variable by IPUMS that assigns each occupation a score based on the median income for that occupation in 1950 U.S. Dollars

Table 2*Linear regression results for age at death and childhood racial designation*

Outcome	Age at Death	
	Race and Cohort Only (1)	Additional Controls (2)
Childhood Race (Reference = White)		
Black	-0.838*** [-0.941, -0.736]	-0.300*** [-0.406, -0.194]
Mulatto	-0.558*** [-0.730, -0.386]	-0.126 [0.300, -0.049]
Finished High School	-	1.043*** [1.011, 1.074]
Occupational Income Score 1940	-	0.003*** [0.002, 0.005]
Birth Year Fixed Effects	yes	yes
State Fixed Effects	no	yes
Intercept	82.970*** [82.879, 83.061]	82.380*** [82.212, 82.549]
N	930,334	917,940
R ²	0.062	0.068
RMSE	7.162	7.138
F-statistic	2,805.211***	937.329***

*Notes: 95% confidence intervals shown.***p < 0.05, **p < 0.01, ***p < 0.001*

Table 3*Education-stratified linear regression results*

	High School Graduates Only (1)	Non-Graduates Only (2)	High School Graduates with Additional Controls (3)	Non-Graduates with Additional Controls (4)
Childhood Race (Reference = White)				
Black	-1.102*** [-1.388, -0.816]	-0.438*** [-0.548, -0.328]	-0.943*** [-1.231, -0.655]	-0.203*** [-0.318, -0.089]
Mulatto	-0.617** [-1.006, -0.227]	-0.264** [-0.456, -0.071]	-0.484* [-0.874, -0.093]	-0.033 [-0.227, 0.162]
Occupational Income Score	-	-	0.006*** [0.004, 0.007]	0.001 [-0.001, 0.003]
Birthyear FE	yes	yes	yes	yes
State FE	no	no	yes	yes
Intercept	83.375*** [83.201, 83.550]	82.796*** [82.689, 82.903]	82.887*** [82.545, 83.228]	82.589*** [82.394, 82.783]
N	367,829	550,111	367,829	550,111
R ²	0.058	0.072	0.058	0.073
RMSE	7.21	7.10	7.20	7.09
F-statistic	1020.830***	1929.564***	321.813***	613.992***

Notes: 95% confidence intervals shown.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4*Regression results using race classification across censuses.*

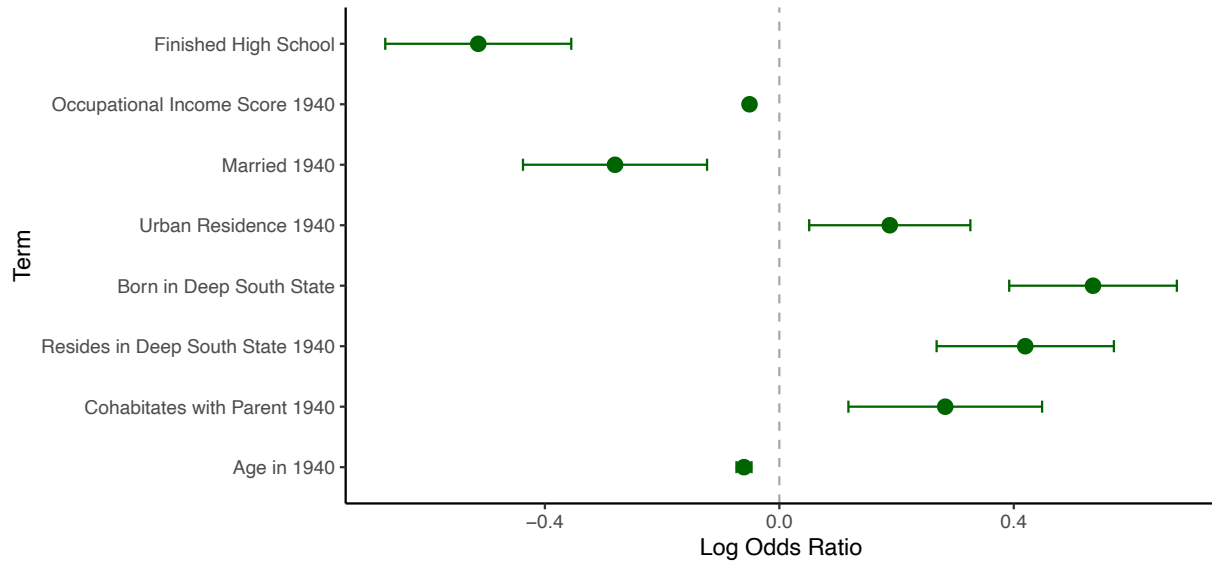
Outcome	Age at Death
	(1)
Race on First Census and 1940 Census (Reference = White-to-White)	
Black-to-Black	-0.870*** [-0.977, -0.763]
Mulatto-to-White	0.090 [-0.355, 0.535]
Mulatto-to-Black	-0.673*** [-0.859, -0.486]
Birth Year Fixed Effects	yes
Intercept	82.969*** [82.877, 83.060]
<hr/>	
N	926,872
R ²	0.062
RMSE	7.16
F-Statistic	2673.283

Notes: 95% confidence intervals shown. Records where a person is recorded as black on one census and white on the other census were not included in the analysis.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1

Variables associated with 1940 race classification



Note: Log odds ratios of variables predicting adult racial classification (white or black) of men originally classified as mulatto. Predictors with positive log odds ratios are associated with black classification; predictors with negative log odds ratios are associated with white classification. All variables except occupational income score and age are binary categorical variables. I define the Deep South as encompassing the states of Alabama, Florida, Georgia, Louisiana, South Carolina, Tennessee, and Texas (this is different from the IPUMS-defined South region). 95% confidence intervals shown.

Table 5*Middle initial availability and agreement rates between matched census records by race across census*

Race in first census and 1940 census	N	Has middle initial in both census records (%)	Middle initial same for records with middle initial available in both census records (%)
White-to-White	902,641	15.3%	86.2%
Black-to-Black	17,520	6.07%	91.9%
Mulatto-to-White	995	12.8%	66.1%
Mulatto-to-Black	5,716	7.00%	87.4%

Note: Records where a person is recorded as black on one census and white on the other census were not included in the analysis.