Black-White Mortality Crossover: New Evidence from Social Security Mortality Records*

Casey F. Breen †

Draft Version: February 21, 2024

Abstract

The Black-White mortality crossover is well-studied demographic paradox. Black Americans experience higher age-specific mortality rates than White Americans throughout most of the life course, but this puzzlingly reverses at advanced ages. The leading explanation for the Black-White mortality crossover centers around selective mortality over the life course. Black Americans who survived higher age-specific mortality risk throughout their life course are highly selected on robustness, and have lower mortality than White Americans in late life. However, skeptics argue the Black-White mortality crossover is simply a data artifact from age misreporting or related data quality issues. We use large-scale linked administrative data (N = 2.3 million) to document the Black-White mortality crossover for cohorts born in the early 20$^{\text{th}}$ century. We find evidence the crossover is not a data artifact and cannot be uncrossed using sociodemographic characteristics alone.

*For helpful feedback and background information, the author would like to thank Dennis M. Feehan, Joshua R. Goldstein, Christopher Muller, and Elizabeth Wrigley-Field. Research reported in this publication was supported by the National Institute of Aging R01AG05894. C.F.B. was supported by National Institute of Aging T32-AG000246.

†University of Oxford. casey.breen@demography.ox.ac.uk.

1 Introduction

The Black-White mortality crossover is a long-standing demographic paradox. The crossover occurs when non-Hispanic Black Americans experience higher age-specific mortality rates than Non-Hispanic Whites Americans until very late in life. At advanced ages, the age-specific mortality rates first converge and then cross over, with Black mortality being lower than White mortality. The crossover has been repeatedly documented in the United States (Sautter et al., 2012; Dupre, Franzese and Parrado, 2006; Masters, 2012; Lynch, Brown and Harmsen, 2003; Hummer, 1996). However, there is little consensus on the explanation for the Black-White mortality crossover: critics have questioned these findings, suggesting that the apparent crossover is simply an artifact of sparse or poor-quality mortality data at the most advanced ages (Preston and Elo, 2006; Lynch, Brown and Harmsen, 2003; Preston et al., 1996; Preston, Elo and Preston, 1999). Others have theorized that the crossover is the product of selective mortality over the life course (Vaupel, Manton and Stallard, 1979; Vaupel and Yashin, 1985; Wrigley-Field, 2014, 2020).

Understanding the Black-White mortality crossover is important for several reasons. First, the mortality crossover has implications for our understanding of inequality at the most advanced ages. Is there really a narrowing of mortality conditions for Black and White Americans among the oldest old? Or is the crossover just a data artifact or a ruse of heterogeneity in susceptibility to mortality? Second, the Black-White mortality crossover is a useful empirical example for developing theoretical frameworks of mortality selection. Finally, insights gained from studying the Black-White mortality crossover can be applied to related research areas, such as mortality compression and deceleration of mortality rates at advanced ages (Lynch, Brown and Harmsen, 2003).

In this study, we use linked administrative mortality data from the CenSoc-DMF (N = 2.3 million) to investigate the Black-White mortality crossover. The unprecedented size of the CenSoc-DMF dataset, along with its rich array of covariates, allows us to empirically assess two of the main explanations for the Black-White crossover. We find a mortality crossover for the male birth cohorts of 1890–1905 at age 85 and a crossover for the male birth cohorts of 1906–1915 at age 90. Our analysis is restricted to men, as surname changes for some women during marriage make linking women between the 1940 Census and the DMF mortality records infeasible. The quality of our mortality data, paired with a sensitivity analysis, allows us to rule out that our observed crossovers are simply an artifact of age misreports or exaggerations. We then stratify for observed heterogeneity to test whether the crossover can be uncrossed using sociodemographic characteristics, finding that the crossover persists across all subgroups. We conclude that unobserved heterogeneity may still be responsible, or there are indeed as-yet unknown protective factors that influence race differentials at older ages in ways that are different than at younger ages.

2 Background

2.1 Past Studies on the Black-White Mortality Crossover

Since its original discovery by Sibley (1930), the Black-White mortality crossover has been repeatedly documented in the United States (Manton, Poss and Wing, 1979; Berkman, Singer and Manton, 1989; Lynch, Brown and Harmsen, 2003; Dupre, Franzese and Parrado, 2006; Sautter et al., 2012; Kestenbaum, 1992; Masters, 2012). The Black-White mortality crossover has also served as a motivating example for a growing body of methodological work on theoretical models of mortality selection (Vaupel and Yashin, 1985; Vaupel, Manton and Stallard, 1979; Wrigley-Field, 2014, 2020). More recently, a handful of empirical studies have investigated the contribution of covariates such as socioeconomic status or religious attendances to the Black-White mortality crossover (Dupre, Franzese and Parrado, 2006; Sautter et al., 2012; Yao and Robert, 2011; Berkman, Singer and Manton, 1989).

Table 1 presents several of the major empirical studies documenting the Black-White crossover. Across studies, the "age of crossover"—the age at which Black age-specific mortality rates first become lower than White age-specific mortality rates—occurs between the ages of 74 and 90, generally centered around 85. However, the age at crossover has been trending upwards over the course of the 20th century (Masters, 2012). In the 1960s, the crossover observed at age 75 for men and age 77 for women (Kestenbaum, 1992). In the 1970s, the age at crossover was observed at ages of 78 for men and 80 for women (Masters, 2012). More recently, the crossover has been observed at ages 88 for men and 87 for women in U.S. lifetables from 2003 (Arias, 2006). This upward trend in the timing of the age of crossover suggests that differential cohort experiences are an important consideration for any study of the Black-White crossover.

Data Source	Age of Crossover	Covariates	Age Verification	Citation
Tennessee Vital Statistics	74			Sibley (1930)
Evans County Study	85 (f); 80 (m)			Wing et al. (1985)
Medicare Enrollment	88 (f); 86 (m)			Kestenbaum (1992)
U.S. Death Certificates	90 (f); 85 (m)		✓	Preston et al. (1996)
Medicare Enrollment	85–86			Parnell and Owens (1999)
Survey on Asset and Health Dynamics Among the Oldest Old	81			Johnson (2000)
Berkeley Mortality Database	79–87		✓	Lynch, Brown and Harmsen (2003)
Medicare Enrollment	80–85			Arias (2006)
Established Populations for Epidemiologic Studies of the Elderly	83 (f); 79 (m)	Religious Attendance		Dupre, Franzese and Parrado (2006)
Americans' Changing Lives study	80	Education, Income, Neighborhood Socioeconomic Disadvantage Index		Yao and Robert (2011)
National Health Interview Survey-Linked Mortality Files	85			Masters (2012)
Established Populations for Epidemiologic Studies of the Elderly	83 (f); 79 (m)			Sautter et al. (2012)
NCHS Multiple Cause-of-Death public-use files	87	Education, Income		Fenelon (2013)
National Longitudinal Mortality Study	85			Şahin and Heiland (2017)

Table 1: Past studies of the Black-White mortality crossover.

2.2 Explanations for the Black-White Crossover

There are three prominent explanations for the Black-White mortality crossover. The evidence to date is not yet seen as conclusive, and population scholars are increasingly seeking explanations for the Black-White mortality crossover. These competing explanations are outlined below.

2.2.1 Data Artifact

One explanation for the Black-White crossover is that there is no crossover at all. Rather, differential age-misreporting or exaggeration, uncounted or unmatched deaths, and other inaccuracies can lead to a spurious crossover. According to this perspective, once these data errors are accounted for, the crossover disappears or is delayed until even more advanced ages (Preston and Elo, 2006; Lynch, Brown and Harmsen, 2003; Preston et al., 1996; Preston, Elo and Preston, 1999).

This perspective was most clearly advanced by Preston et al. (1996), who linked death certificates to both decennial census records (1900, 1910, and 1920) and the Social Security Death Master File (DMF). This linkage exercise demonstrated that misreporting was common; over 50% of Black women decedents had disagreement between the ages of death on their death certificate and their Social Security record. Upon correcting for misreporting in these death rates for Black Americans, the crossover disappeared. As further evidence of age misreporting, Preston and Elo (2006) in a follow-up study demonstrated that the age-specific mortality rates for Black Americans above 85 were lower than the age-specific mortality rates in the lowest-mortality countries.

2.2.2 Age-As-A-Leveler

The "naive" theoretical explanation for the Black-White mortality crossover is that for the oldest-old, mortality conditions converge for Black and White Americans. According to this *age-as-a-leveler* hypothesis, older adults are increasingly separated from the unequal social institutions that contribute to racial health disparities, such as the education system, the labor market and the criminal justice system. The departure from these stressors of daily living may cause mortality rates to converge in later life (Kim and Miech, 2009). Further, increased availability of a social safety net in later life, including Medicare and Social Security, and stronger kin and support networks, could cause age-specific mortality rates to converge in the oldest ages.

In this sense, old age acts as a "leveler" and causes a convergence in age-specific mortality rates; real racial disadvantage attenuates at the most advanced age. However, it is unclear why such attenuation of disadvantage at the most advanced ages would cause a crossover, rather than simply a convergence. Further, this hypothesis is at odds with a large body of research documenting racial inequality in the U.S. (Bryan L. Sykes and Michelle Maroto, 2016; Alexander, 2010; Riddle and Sinclair, 2019; Perry and Morris, 2014).

2.2.3 Heterogeneity in Frailty

The most famous explanation for the mortality crossover comes from theoretical models of mortality selection. Mortality selection models begin with the premise that people vary systematically in mortality risk. In this frailty modeling tradition, as a cohort ages, it becomes increasingly composed of robust individuals. This mortality selection can occur unequally across population subgroups, and has been hypothesized to explain mortality crossovers, mortality deceleration, and mortality compression (Lynch, Brown and Harmsen, 2003; Wrigley-Field, 2014).

In the case of the Black-White crossover, Black Americans who faced higher mortality risks in early and midlife will be composed of a greater proportion of robust individuals in later life, resulting in their age-specific mortality rates becoming lower than those of White Americans, who faced lower mortality risks earlier in their life course. In other words, the Black Americans who survive to the most advanced ages are more highly selected for robustness than their White counterparts, and will have lower mortality at advanced ages (Wrigley-Field, 2020; Vaupel and Yashin, 1985).

Skeptics of this heterogeneity in frailty explanation point out that poor health conditions in early life can "scar" survivors, leading to higher mortality in later life (Preston and Elo, 2006). The limited number of empirical investigations have suggested that the dominant direction of mortality conditions at different points in the life course is positively: higher mortality risk in early life is associated with higher mortality later in the life course (Finch and Crimmins, 2004; Janssen et al., 2004; Preston, 1970).

This study has two specific aims. First, we establish the mortality crossover as real, not a data artifact. Second, we provide empirical evidence that the crossover cannot be uncrossed using sociodemographic characteristics alone. The remainder of the paper proceeds as follows. In the Section 3, we describe the complete count census data and mortality records used in our analysis. We then describe our methods for mortality estimation in the absence of denominators in Section 4. In Section 5 and Section 6, we present and interpret our findings and discuss their implications for our understanding of mortality selection.

3 Data

This study uses complete count 1940 Census data, mortality records from the Social Security Death Master File (DMF), and record linkage techniques to construct a large-scale dataset with rich covariates and mortality outcomes. This dataset, termed the CenSoc-DMF (Goldstein et al., 2021), links the complete count 1940 Census (Ruggles et al., 2020) to the DMF. The DMF is a collection of over 83 million death records reported to the Social Security Administration, with nearly complete mortality coverage between 1975–2005 (Alexander, 2018; Hill, 2001). However, the DMF does not contain any socioeconomic or demographic variables. To obtain individual-level covariates, we link the DMF mortality records to 1940 Census records. The resulting matched file includes only men, as surname changes due to marriage for some women make the systematic linkage of women infeasible.

We link individual records in the complete count 1940 Census to the DMF using first name, last name, and year of birth using the ABE exact match record linkage algorithm (Abramitzky, Boustan and Eriksson, 2012, 2014; Abramitzky and Boustan, 2017; Abramitzky et al., 2021). To reduce false matches, we restrict to matches where names are unique within and across datasets for a $\pm$ 2 year window. This approach prioritizes minimizing the number of false matches over maximizing the overall match rate; this minimizes the amount of systematic bias introduced by false matches (Ruggles, Fitch and Roberts, 2018).

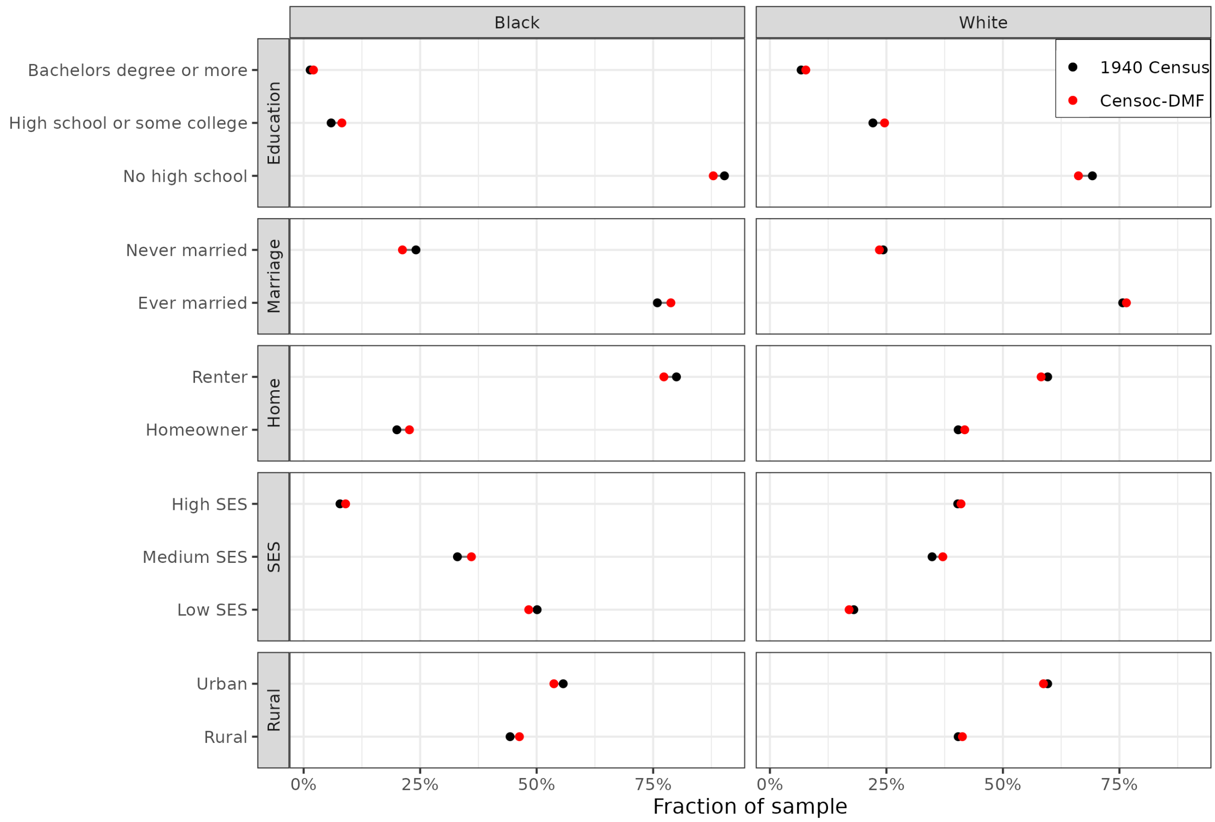

Figure 1: Each facet shows the composition of the CenSoc-DMF (Red) and the complete count 1940 Census (black) for a given covariate for Black and White matches. The matched sample has slightly higher socioeconomic status than the general population.

3.1 Representativeness of Matches

Our mortality-adjusted match rate is approximately 20% (Breen and Osborne, 2022). To demonstrate that our matched sample is representative of the general population within racial groups, we compare the composition of our matched sample to the general population in the 1940 Census. Figure 1 shows our matched sample is broadly representative of the general population within racial groups, except with slightly higher socioeconomic status.

3.2 Reliability of DMF

A key consideration for our study is the reliability of the DMF mortality records. The DMF is extracted from the Social Security Numident, and contains over 75 million death records. The death coverage between 1975–2005 is nearly complete, containing approximately 95%

coverage for deaths occurring after the age of 65 (Hill, 2001; Alexander, 2018). Death coverage rates drop after after 2005, and the DMF has substantial coverage gaps beginning in 2011 (Maynard, 2019). Our analysis is restricted to deaths occurring in our mortality observation window of 1975–2005.

The DMF does not explicitly include information on age of death. Rather, the DMF contains information on date of birth and date of death from which age of death can be imputed (Preston et al., 1996). Therefore, to assess the reliability of the imputed age of death, we need to investigate the quality of the reported date of birth and date of death.

Dates of death are directly reported to the Social Security Administration from a funeral director or a family member. These reports are generally made directly following a death, minimizing the likelihood of misreporting. The date of death in the DMF almost always exactly matches the date of death in the corresponding death certificates (Hill, Preston and Rosenwaike, 2000).

Information on date of birth is submitted personally by the decedent in conjunction with a benefit claim. The Social Security Administration closely tracks age to determine eligibility for benefits. Age verification is a required condition for entitlement to benefits, and stringent tests were put in place in 1965. The focal cohorts of this study would have become eligible for Social Security benefits after these age verification procedures were put in place.

To empirically assess the reliability of the date of birth information in the DMF, we look at heaping on year of birth. Heaping, a common indicator of data quality, is the systematic misstatement of ages or dates to round or terminal ages (e.g., end in "0" or "5.") We find minimal date heaping on year of birth, as shown in Figure 2. However, there is slightly higher heaping for Black Americans than White Americans. To investigate whether this age heaping has any affect on our observed crossover, we conduct a sensitivity analysis by dropping years of birth that end in terminal ages and re-estimating the observed crossover.

The nature of our sample provides additional reassurance that the reported age of birth is accurate. For an individual to be successfully matched and included in our sample, their reported age in the 1940 Census must correspond to ± 2 years of their year of birth reported in the DMF. Therefore, mortality records where the year of birth is misreported by over two years will be excluded from our sample. This is similar to the validation approach taken

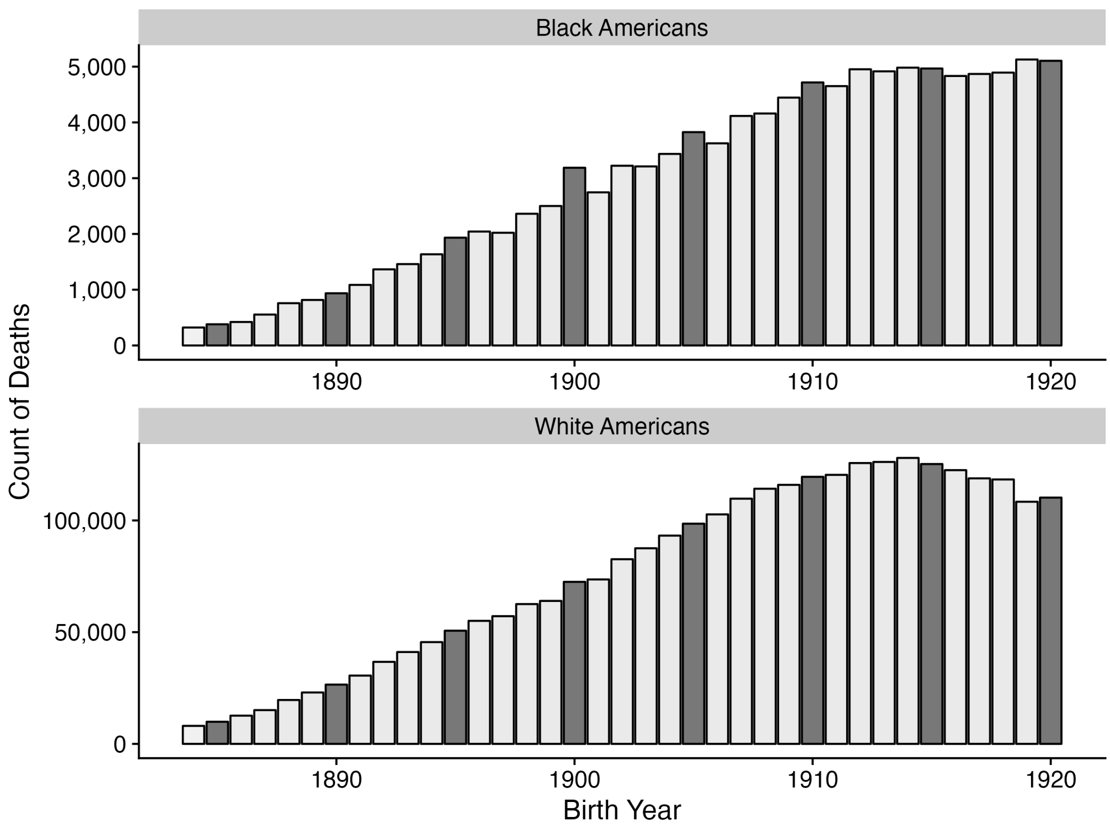

Figure 2: Highlighted years (dark grey) show very slight amounts of age-heaping on the terminal digits “0” or “5,” suggesting the DMF has minimal misreporting of year of birth.

by Hill, Preston and Rosenwaike (2000); Preston et al. (1996), and gives an additional level of reassurance that the reported birth year is accurate.

4 Methods

4.1 Estimating Mortality Rates

The CenSoc-DMF dataset only includes deaths for the left and right ("doubly") truncated window of 1975 to 2005. Further, the CenSoc-DMF does not include any measure of survivorship, as we have no way of determining whether an individual observed in the 1940 Census died outside our observation window window or was not successfully matched to their death record. The absence of any measure of a denominator precludes conventional occurrence-exposure methods for estimating mortality rates (Alexander, 2018).

To overcome this, we use two different methods to estimate mortality rates in the absence of denominators. First, for the earlier cohorts of 1890–1905, we use the reverse survival method to estimate mortality rates. This approach assumes that all persons in the cohort have died by the end of our mortality observation window in 2005. Specifically, we estimate the total number of survivors at a given age by summing up all the deaths occurring above that age, and then estimating the age-specific mortality rates using the age-specific ratios of deaths to survivors. This method is only appropriate for the cohorts born before 1905, for which only a few survivors to age 100 will die after 2005.

Second, for the later-born cohorts of 1906-1915, those that we cannot assume are extinct by 2005, we assume the distribution of deaths within a cohort follows a Gompertz distribution and use maximum likelihood estimation methods to estimate the parameters of this distribution (Goldstein et al., 2023; Gompertz, 1825). Specifically, the hazard of dying at age x is:

$$h_0(x) = ae^{bx} \quad (1)$$

where a is a background level of mortality at age x, b is the rate of mortality increase with age, and $h(x)$ describes the hazard schedule. This approach allows us to estimate age-specific mortality rates for both ages where we did and did not observe deaths.

4.2 Stratifying on Observed Dimensions of Heterogeneity

The classical mortality selection model used to explain the crossover is unidimensional. That is, all heterogeneity in susceptibility to mortality is captured in a single parameter ("frailty"). A growing body of empirical research on the Black-White mortality crossover has used individual-level covariates to study the observed dimensions of heterogeneity that constitute frailty. Borrowing logic from unidimensional mortality selection model, these studies investigated how controlling for some piece of frailty changes the age of crossover (Sautter et al., 2012; Dupre, Franzese and Parrado, 2006). Yet theoretical advances have demonstrated that the unidimensional mortality selection model is not an appropriate starting point for empirical work. When there is both observed and unobserved heterogeneity, stratifying on observed heterogeneity can cause the age at crossover to either move up or down (Wrigley-Field, 2020). In other words, the age at crossover will always change when some factor related to both race and mortality is controlled for.

One important exception occurs when an observed dimension of heterogeneity constitutes a large portion of the overall heterogeneity. In this setting, if the crossover is caused by heterogeneity in frailty, stratifying on a covariate that represents over 50% of total frailty will uncross the crossover (Wrigley-Field, 2020). For empirical researchers, this implies that combining many covariates into a single risk measure is a promising strategy for examining the role of observed heterogeneity in explaining the crossover.

To investigate the role that observed heterogeneity plays on the mortality crossover, we use socioeconomic covariates available in the 1940 Census. First, we investigate the crossover in six distinct subgroups: individuals with high education (more than 8 years), individuals with low education (less than 8 years), individuals with high income (above the median income), individuals with low income (below the median income), homeowners, and renters. On each subgroup, we estimate age-specific mortality rates using the reverse survival method. We then combine these covariates into a single risk score and investigate the crossover in subgroups defined by risk. Together, these analyses allows us to investigate whether the crossover still persists when we stratify on major pieces of frailty.

5 Results

We first analyze the Black-White mortality crossover using both the reverse survival method and our parametric Gompertz approach. Next, we present results on observed mortality selection. Finally, we examine whether our observed heterogeneity can help explain the Black-White mortality crossover.

5.1 Black-White Mortality Crossover

We first examine the Black-White crossover for the pooled birth cohorts of 1890-1905. Figure 3a shows a clear mortality crossover at age 86, consistent with past findings. For this analysis, we estimated age-specific mortality rates using the reverse survival method. Because our mortality data showed very slight heaping on year of birth, as a sensitivity analysis, we recalculate our age-specific mortality rates excluding birth years with potential age heaping: 1890, 1895, 1900, and 1905. Figure 3b shows that the crossover persists, suggesting that low quality mortality data is not responsible for the crossover.

For the cohorts of 1905-1915, which are not extinct by the end of our mortality observation window in 2005, we fit a parametric Gompertz model to calculate age-specific mortality rates (Goldstein et al., 2023). We perform maximum likelihood estimation for the Black and White groups separately. Figure 4 shows a mortality crossover at age 90, slightly higher than our observed age at crossover for the cohorts of 1890–1905. A higher age at crossover for later birth cohorts is consistent with past studies (Masters, 2012).

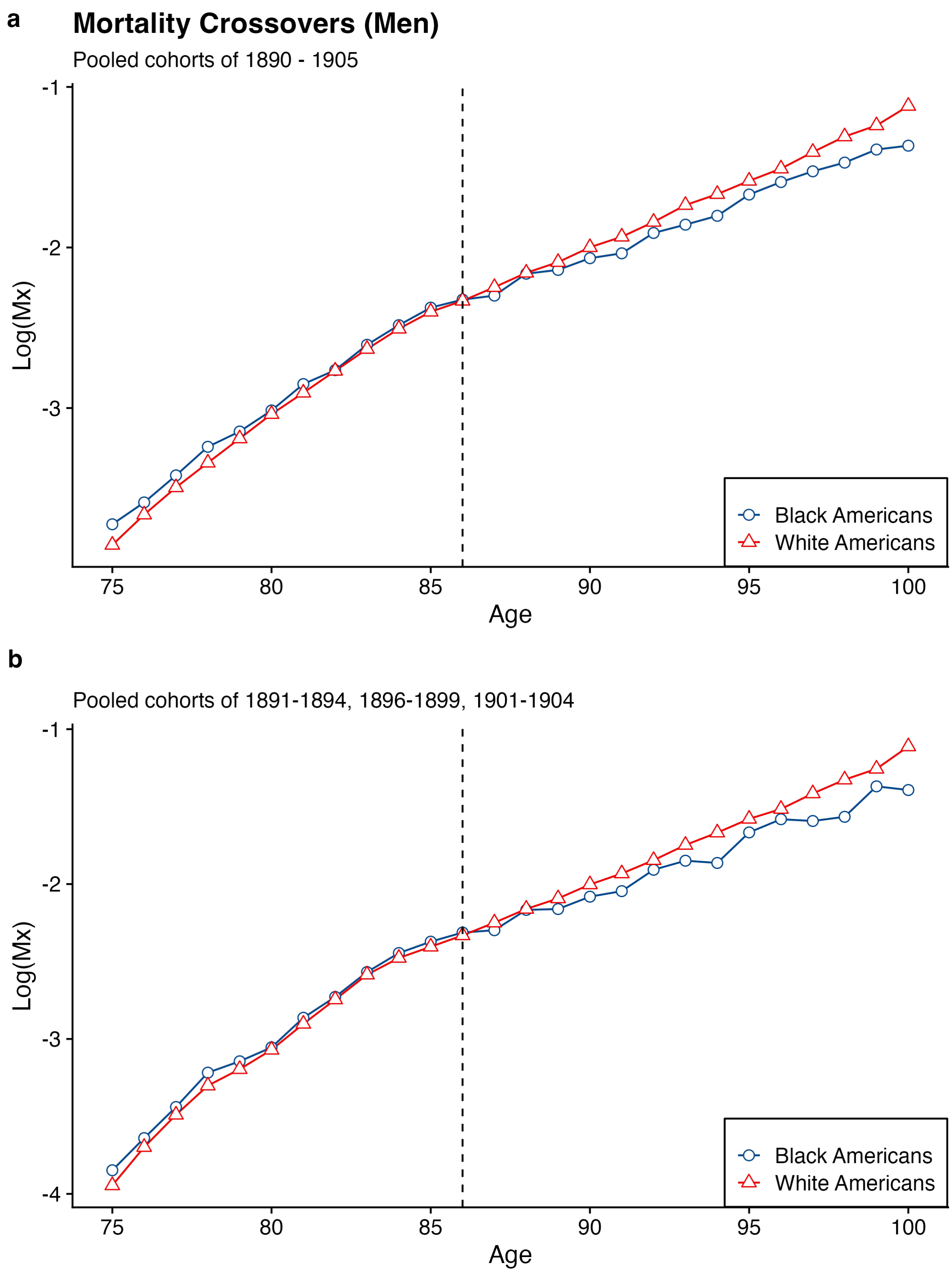

Figure 3: Panel (a) shows the Black-White mortality crossover for the cohorts of 1890-1905. Panel (b) shows the mortality crossover dropping the cohorts of 1890, 1895, 1900, and 1905, where we observed slight but detectable age heaping. The mortality rates were estimated using the reverse survival method.

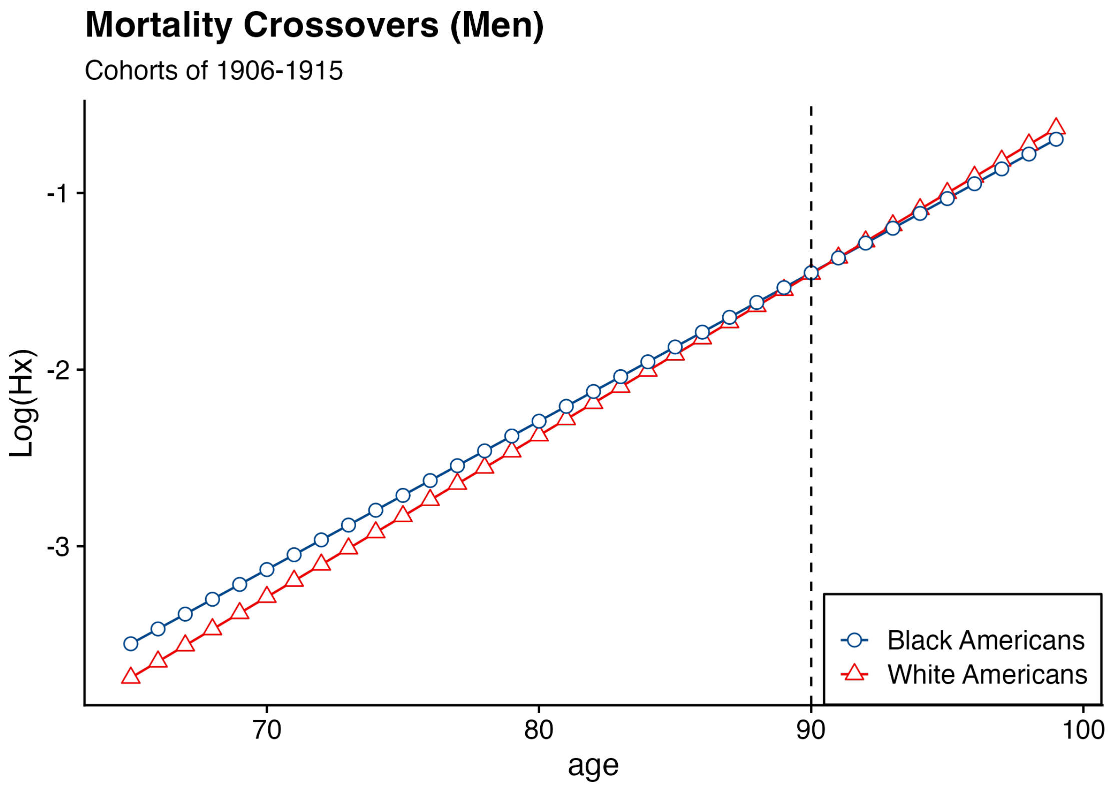

Figure 4: Black White mortality crossovers for cohorts of 1905-1915.

5.2 Observed Mortality Selection

To investigate mortality selection, we track how the characteristics of survivors change as a cohort ages and members of the cohort die off. We focus on how the composition of the cohorts of 1909–1911 changes with respect to employment, educational attainment, socioeconomic status score, wage and salary income, homeownership status, and residing in the south. We interpret an increase in a dimension of socioeconomic status as a cohort ages to be evidence of selective mortality: more frail individuals are dying off at earlier ages.

As shown in Figure 5, we do observe selective mortality, which is more pronounced for White Americans than Black Americans. For instance, members of the cohort of 1909–1911 who survived to age 65 have approximately 10 years of education, while members of the cohort who survived to age 90 have approximately 10.6 years of education. The difference is more slight for the cohort of Black Americans: survivors at age 65 had 6.6 years of education, and survivors at age 90 had 6.7 years of education. Across all of the covariates tested, we find that the the surviving members of a cohort becoming more increasingly advantaged as the cohort ages.

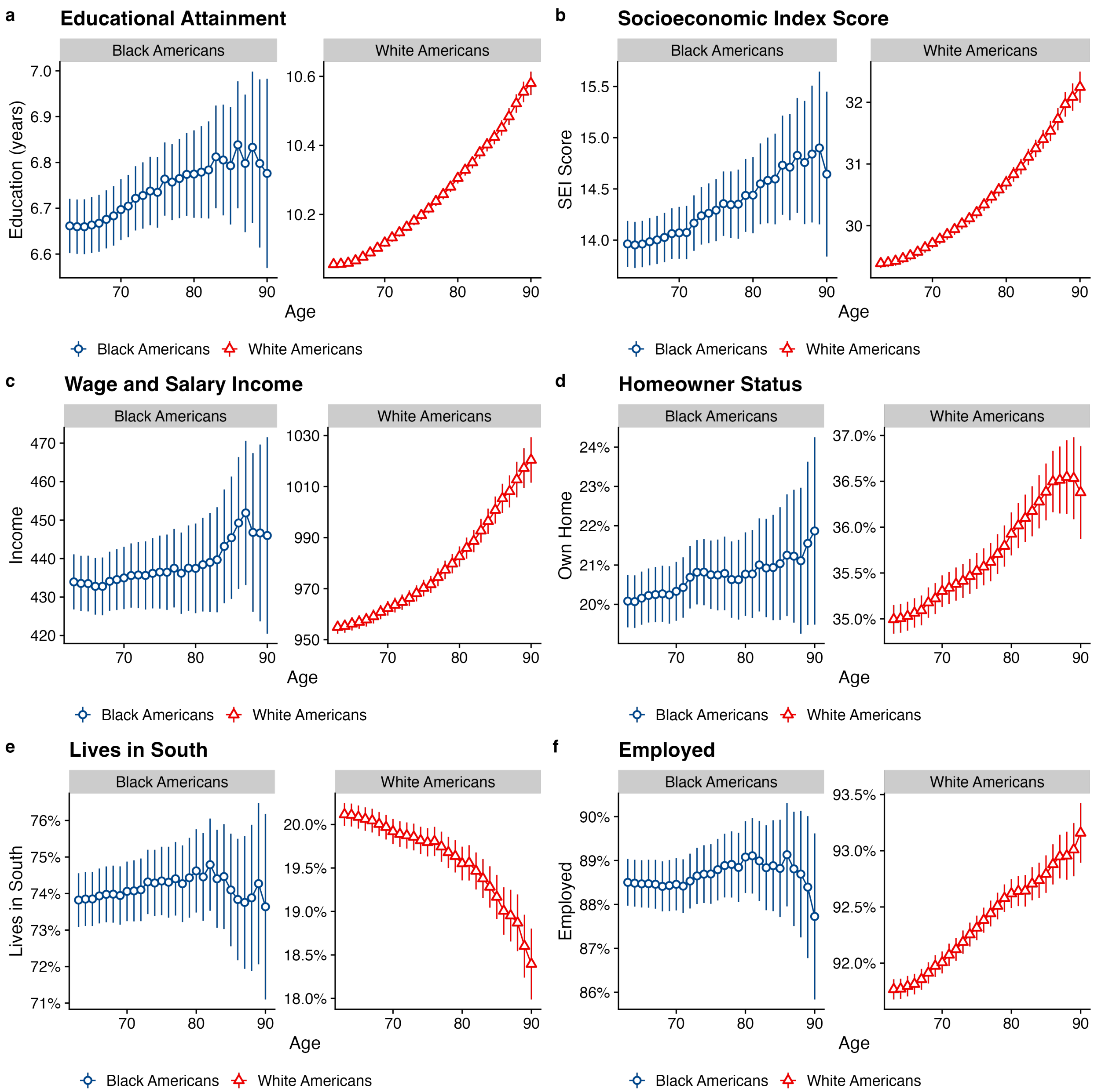

Figure 5: Changing composition of the survivors. We see only modest evidence of selection. Error bands show 95% uncertainty intervals.

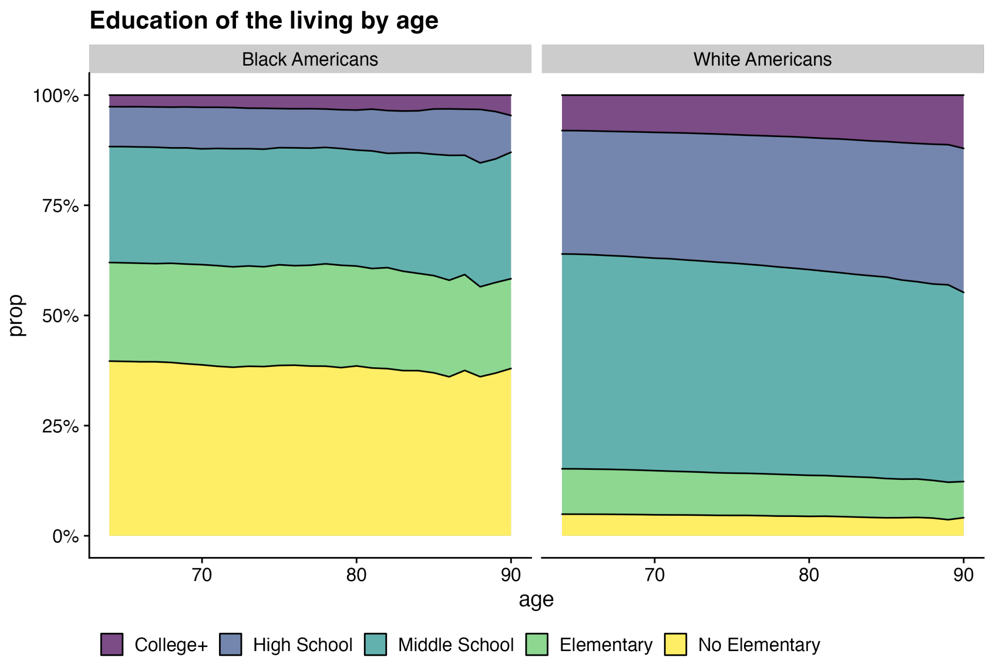

Figure 6: Changing educational composition of the survivors.

5.3 The Surprising Non-Effect of Observed Heterogeneity on the Mortality Crossover

Next, we investigate the effect of observed heterogeneity on the morality crossover. We split our 1890-1905 birth cohort sample into different population subgroups defined by education, homeownership, and wage and salary income. Figure 7 shows the result of this analysis: the crossover persists across all subgroups.

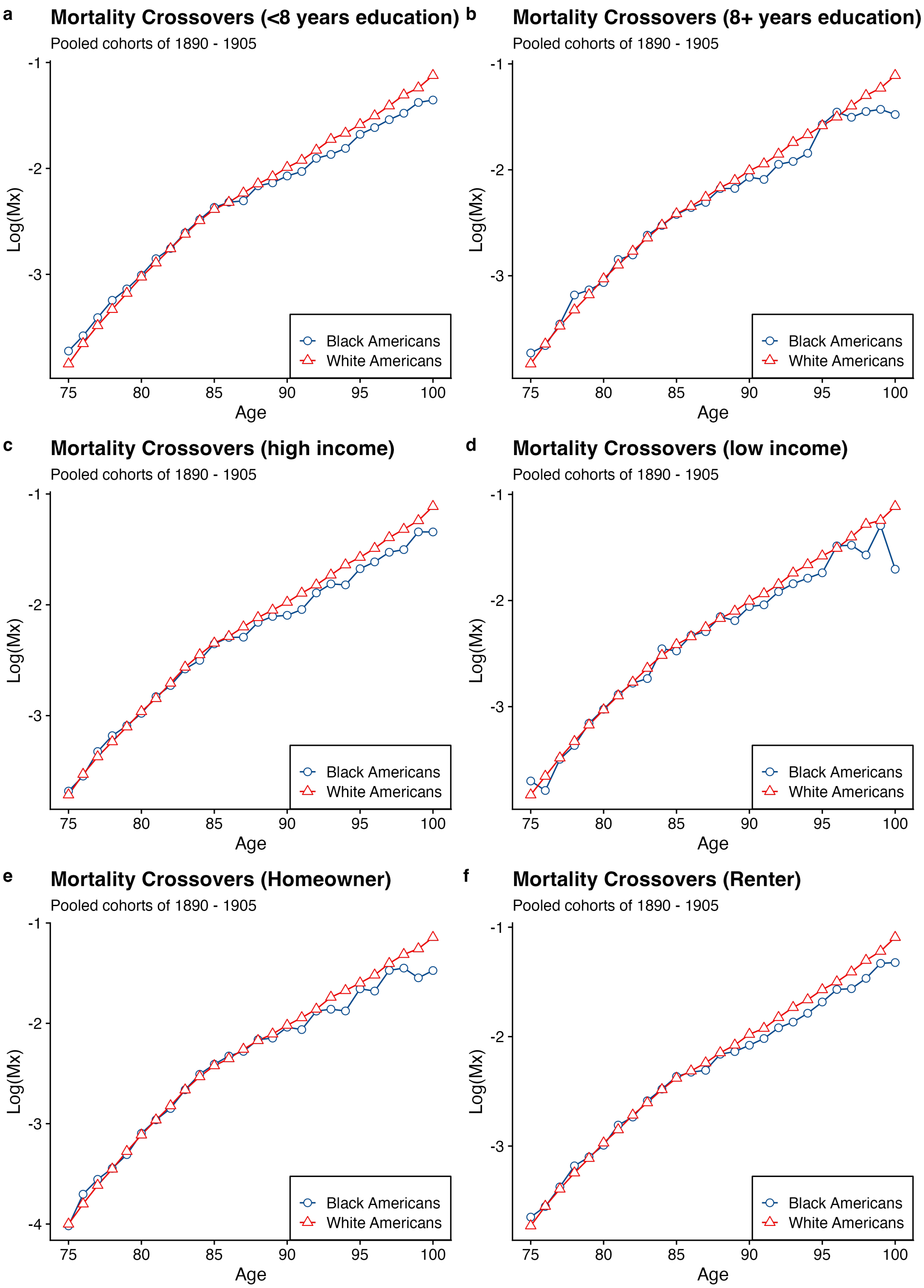

Figure 7: Black White mortality crossover for different subgroups defined by socioeconomic status.

Next, we follow the advice outlined in Wrigley-Field (2020) and investigate the mortality crossover stratified by risk scores. To construct the risk score, we aggregate together the following covariates into a single score: education, wage and salary income, socioeconomic index, marital status, employment, living in the South, and owning a home. The motivation for constructing this risk score is to capture as much of the heterogeneity in frailty as possible in one score. To estimate the risk score, we fit linear regressions of the form:

$$\text{death\_age}_i = educ_i + income_i + homeowner_i + marital\_status_i + southern_i + \epsilon_i. \quad (2)$$

Figure 8a presents the age-specific mortality rates for Black and White men within risk group. We see the crossover persists in all different risk groups. Figure 8b plots the difference in log hazards, again finding a clear crossover for all three groups.

[0]We first fit this model on our analytic sample, those born between 1890-1905. To avoid potentially over-fitting, we also fit the model on the out-of-sample cohort of 1906. Results from both models provided highly comparable predictions of risk score.

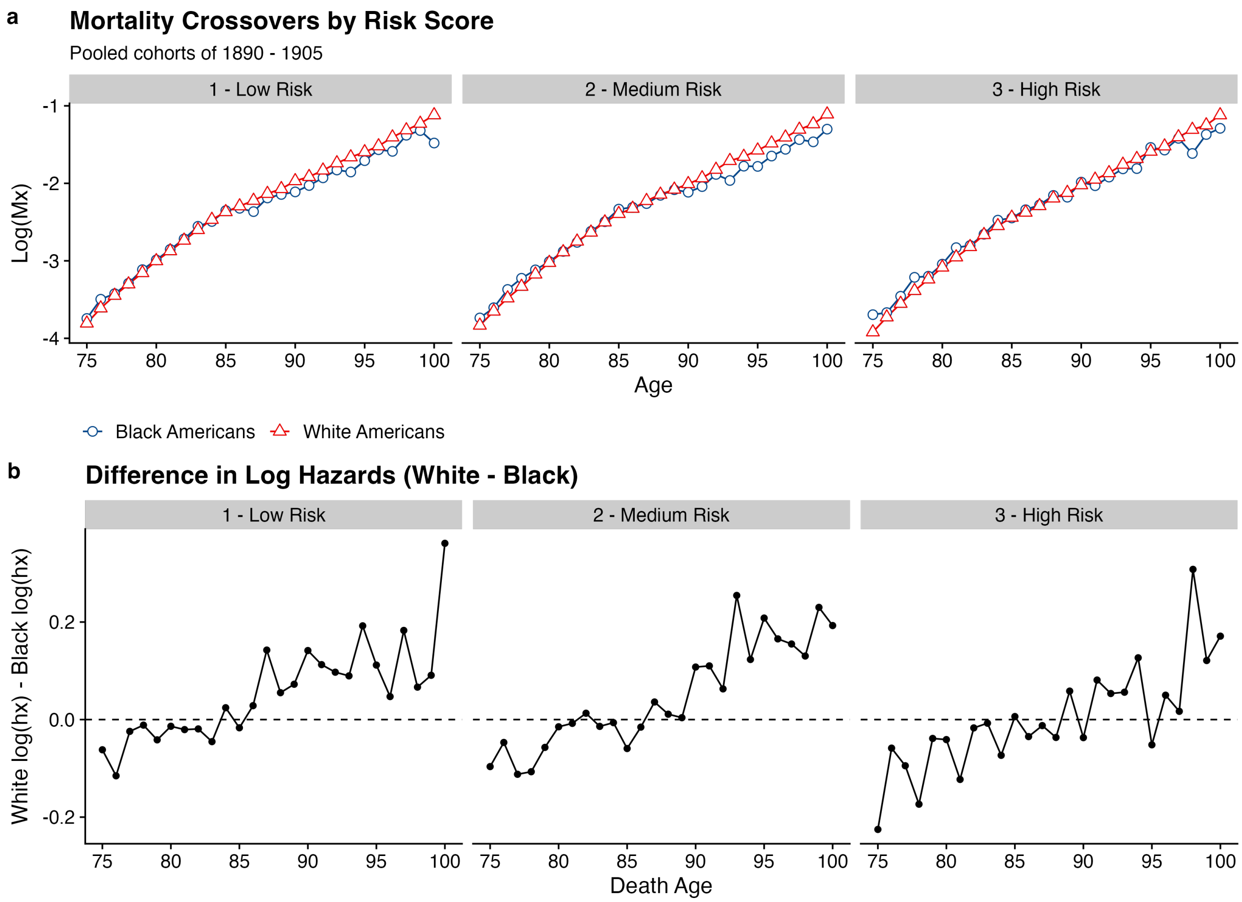

Figure 8: Black White mortality crossover by risk score.

6 Discussion

This study uses mortality records from the Death Master File (DMF) linked to the 1940 Census to investigate the Black-White mortality crossover. We find a clear mortality crossover at age 85 for men in the birth cohorts of 1890–1905 using reverse survival methods to estimate age-specific mortality rates. Using a Gompertz parametric maximum likelihood approach, we find a mortality crossover at age 90 for the birth cohorts of 1906–1915. Given the reliability of the DMF mortality data, we interpret this as evidence that the Black-White mortality crossover is not simply an artifact of sparse data or age misreporting: the crossover persists even when we restrict the sample to the highest-quality mortality data.

Using individual-level characteristics from the 1940 Census, we investigate observable mortality selection. We find clear evidence of selective mortality: as a cohorts ages, the survivors have increasingly higher educational attainment, rates of homeownership, rates of employment (in 1940), and wage and salary income. However, the observable selection is relatively modest, and is more pronounced for White Americans than Black Americans. The lack of observable mortality selection for Black Americans is perhaps attributable to the weaker correlation between covariates, such as education or income, and mortality risk for Black Americans (Card and Krueger, 1992).

Our investigation of the Black-White mortality crossover for subgroups defined by socioeconomic characteristics indicated a clear crossover in every subgroup. Additionally, the crossover persisted when we stratified on a risk score that aggregated many mortality covariates. This suggests that stratifying on observed socioeconomic dimensions of heterogeneity does not explain the crossover. There are two potential explanations for this finding. First, it is possible that sociodemographic characteristics alone simply do not capture enough of the heterogeneity in frailty to really uncross the crossover. Second, it is possible that the crossover is not driven by heterogeneity in frailty at all; rather, there is actually some true narrowing of inequality at the most advanced ages.

Taken together, our results suggest that the mortality crossover is real and not an artifact of measurement or data errors. Our data allows us to study the mortality experience of real cohorts, not the synthetic period measures commonly used to study the crossover. However, our study cannot make definitive about the theoretical explanations for the crossover. While our study found that stratifying on observed dimensions of frailty such as educational attainment or homeownership does not explain the crossover, it is possible that we are simply not capturing enough of the heterogeneity in frailty to uncross the crossover.

There are several limitations and avenues for future research. First, we only observe mortality window of 1975–2005, so our analyses are restricted to birth cohorts that would be experiencing a crossover in our mortality observation window. Second, it is possible that the sociodemographic characteristics we observe only constitute a very small piece of frailty and therefore have limited utility for explaining the crossover. Future research could test whether covariates that capture more of the heterogeneity in frailty, such as biomarkers, anthropometric measures (weight, height), and direct measurement of subjective and objective health on the crossover. It is important to acknowledge that the present study benefits from an exceptionally large sample size, making it potentially challenging for other studies to achieve comparable levels of precision. Third, while we find little evidence of age misstatement or exaggeration, it is possible there remain undetected age misreports in the DMF. Finally, this analysis is limited in scope to men. Broadening this study to include women is necessary to make complete claims about health and longevity disparities in the most advanced ages.

References

Abramitzky, Ran and Leah Boustan. 2017. "Immigration in American Economic History." *Journal of Economic Literature* 55(4):1311–1345.

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.

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2012. "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review* 102(5):1832–1856.

Abramitzky, Ran, Leah Platt Boustan and Katherine Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122(3):467–506.

Alexander, Michelle. 2010. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New Press.

Alexander, Monica. 2018. Deaths without Denominators: Using a Matched Dataset to Study Mortality Patterns in the United States. Preprint SocArXiv.

Arias, Elizabeth. 2006. "National Vital Statistics Reports, Vol. 54, No. 14 (4/19/2006).".

Berkman, Lisa, Burton Singer and Kenneth Manton. 1989. "Black/White Differences in Health Status and Mortality Among the Elderly." *Demography* 26(4):661–678.

Breen, Casey and Maria Osborne. 2022. An Assessment of CenSoc Match Quality. Preprint SocArXiv.

Bryan L. Sykes and Michelle Maroto. 2016. "A Wealth of Inequalities: Mass Incarceration, Employment, and Racial Disparities in U.S. Household Wealth, 1996 to 2011." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 2(6):129.

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

Dupre, Matthew E., Alexis T. Franzese and Emilio A. Parrado. 2006. "Religious Attendance and Mortality: Implications for the Black-White Mortality Crossover." *Demography* 43(1):141–164.

Fenelon, Andrew. 2013. "An Examination of Black/White Differences in the Rate of Age-Related Mortality Increase." *Demographic Research* 29:441–472.

Finch, Caleb E. and Eileen M. Crimmins. 2004. "Inflammatory Exposure and Historical Changes in Human Life-Spans." *Science* 305(5691):1736–1739.

Goldstein, J. R., M. Alexander, C. Breen, A. Miranda González, F. Menares, M. Osborne, M. Snyder and U. Yildirim. 2021. "Censoc Project." *CenSoc Mortality File: Version 2.0. Berkeley: University of California* .

Goldstein, Joshua R, Maria Osborne, Casey F Breen and Serge Atherwood. 2023. "Mortality Modeling of Partially Observed Cohorts Using Administrative Death Records." *Soc Arxiv.* p. 33.

Gompertz, Benjamin. 1825. "On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies." *Philosophical Transactions of the Royal Society of London* .

Hill, Mark E. 2001. "The Social Security Administration's Death Master File: The Completeness of Death Reporting at Older Ages." *Social Security Bulletin* 64(1).

Hill, Mark E., Samuel H. Preston and Ira Rosenwaike. 2000. "Age Reporting among White Americans Aged 85+: Results of a Record Linkage Study." *Demography* 37(2):175–186.

Hummer, Robert A. 1996. "Black-White Differences in Health and Mortality:." *The Sociological Quarterly* 37(1):105–125.

Janssen, I., L. Krabbendam, M. Bak, M. Hanssen, W. Vollebergh, R. de Graaf and J. van Os. 2004. "Childhood Abuse as a Risk Factor for Psychotic Experiences." *Acta Psychiatrica Scandinavica* 109(1):38–45.

Johnson, N. E. 2000. "The Racial Crossover in Comorbidity, Disability, and Mortality." *Demography* 37(3):267–283.

Kestenbaum, B. 1992. "A Description of the Extreme Aged Population Based on Improved Medicare Enrollment Data." *Demography* 29(4):565–580.

Kim, JinYoung and Richard Miech. 2009. "The Black-White Difference in Age Trajectories of Functional Health over the Life Course." *Social science & medicine (1982)* 68(4):717.

Lynch, Scott M., J. Scott Brown and Katherine G. Harmsen. 2003. "Black-White Differences in Mortality Compression and Deceleration and the Mortality Crossover Reconsidered." *Research on Aging* 25(5):456–483.

Manton, K. G., S. S. Poss and S. Wing. 1979. "The Black/White Mortality Crossover: Investigation from the Perspective of the Components of Aging." *The Gerontologist* 19(3):291–300.

Masters, Ryan K. 2012. "Uncrossing the U.S. Black-White Mortality Crossover: The Role of Cohort Forces in Life Course Mortality Risk." *Demography* 49(3):773–796.

Maynard, Charles. 2019. "The Incompleteness of the Social Security Death Master File." *JAMA Cardiology* 4(8):831.

Parnell, Allan M. and Cynthia R. Owens. 1999. "Evaluation of U.S. Mortality Patterns at Old Ages Using the Medicare Enrollment Data Base." *Demographic Research* 1:2.

Perry, Brea L. and Edward W. Morris. 2014. "Suspending Progress: Collateral Consequences of Exclusionary Punishment in Public Schools." *American Sociological Review* 79(6):1067–1087.

Preston, Samuel H. 1970. "An International Comparison of Excessive Adult Mortality." *Population Studies* 24(1):5–20.

Preston, Samuel H. and Irma T. Elo. 2006. "Black Mortality at Very Old Ages in Official US Life Tables: A Skeptical Appraisal." *Population and Development Review* 32(3):557–566.

Preston, Samuel H., Irma T. Elo, Ira Rosenwaike and Mark Hill. 1996. "African-American Mortality at Older Ages: Results of a Matching Study." *Demography* 33(2):193.

Preston, Samuel H., Irma T. Elo and Samuel H. Preston. 1999. "Effects of Age Misreporting on Mortality Estimates at Older Ages." *Population Studies* 53(2):165–177.

Riddle, Travis and Stacey Sinclair. 2019. "Racial Disparities in School-Based Disciplinary Actions Are Associated with County-Level Rates of Racial Bias." *Proceedings of the National Academy of Sciences* 116(17):8255–8260.

Ruggles, Steven, Catherine A. Fitch and Evan Roberts. 2018. "Historical Census Record Linkage." *Annual Review of Sociology* 44(1):19–37.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Mathew Sobek. 2020. "IPUMS USA: Version 10.0 [Dataset]." *Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V10.0.* .

Şahin, Duygu Başaran and Frank W. Heiland. 2017. Black-White Mortality Differentials at Old-Age: New Evidence from the National Longitudinal Mortality Study*. In *Applied Demography and Public Health in the 21st Century*, ed. M. Nazrul Hoque, Beverly Pecotte and Mary A. McGehee. Applied Demography Series Cham: Springer International Publishing pp. 141–162.

Sautter, Jessica M., Patricia A. Thomas, Matthew E. Dupre and Linda K. George. 2012. "Socioeconomic Status and the Black–White Mortality Crossover." *American Journal of Public Health* 102(8):1566–1571.

Sibley, Elbridge. 1930. *Differential Mortality in Tennessee, 1917-1928.* Fisk University Press.

Vaupel, J. W. and A. I. Yashin. 1985. "Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics." *The American Statistician* 39(3):176–185.

Vaupel, James W., Kenneth G. Manton and Eric Stallard. 1979. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16(3):439–454.

Wing, Steve, Kenneth G. Manton, Eric Stallard, Curtis G. Hames and H. A. Tryoler. 1985. "The Black/White Mortality Crossover: Investigation in a Community-Based Study1." *Journal of Gerontology* 40(1):78–84.

Wrigley-Field, Elizabeth. 2014. "Mortality Deceleration and Mortality Selection: Three Unexpected Implications of a Simple Model." *Demography* 51(1):51–71.

Wrigley-Field, Elizabeth. 2020. "Multidimensional Mortality Selection: Why Individual Dimensions of Frailty Don't Act Like Frailty." *Demography* 57(2):747–777.

Yao, Li and Stephanie A. Robert. 2011. "Examining the Racial Crossover in Mortality between African American and White Older Adults: A Multilevel Survival Analysis of Race, Individual Socioeconomic Status, and Neighborhood Socioeconomic Context." *Journal of Aging Research* 2011:1–8.