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

Cornell Population Center

Casey F. Breen

University of Texas at Austin

October 17, 2025

Black-White differences in mortality in the U.S.

Black-White differences in mortality in the United States are:

- ▶ Huge

- ▶ Historic

- ▶ Ongoing
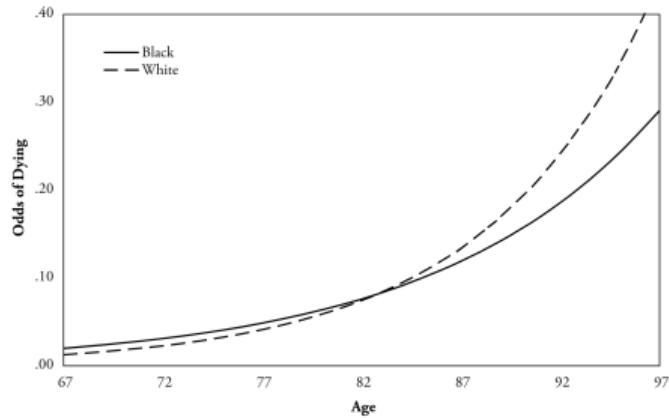
PNAS. Wrigley-Field 2020.

Paradox: Black-White mortality crossover

▶ Yet — among oldest-old, Black
mortality is lower than White
mortality (!)

Paradox: Black-White mortality crossover

▶ Yet — among oldest-old, Black mortality is lower than White mortality (!)

▶ Black-White mortality crossover is a well-studied demographic **paradox**

Dupre 2006. *Demography.*

Why understanding Black-White crossover is important

1. Fundamental to our understanding of **inequality** over the life course

Why understanding Black-White crossover is important

1. Fundamental to our understanding of **inequality** over the life course

2. Oldest-old is fastest growing age segment, potential social policy implications
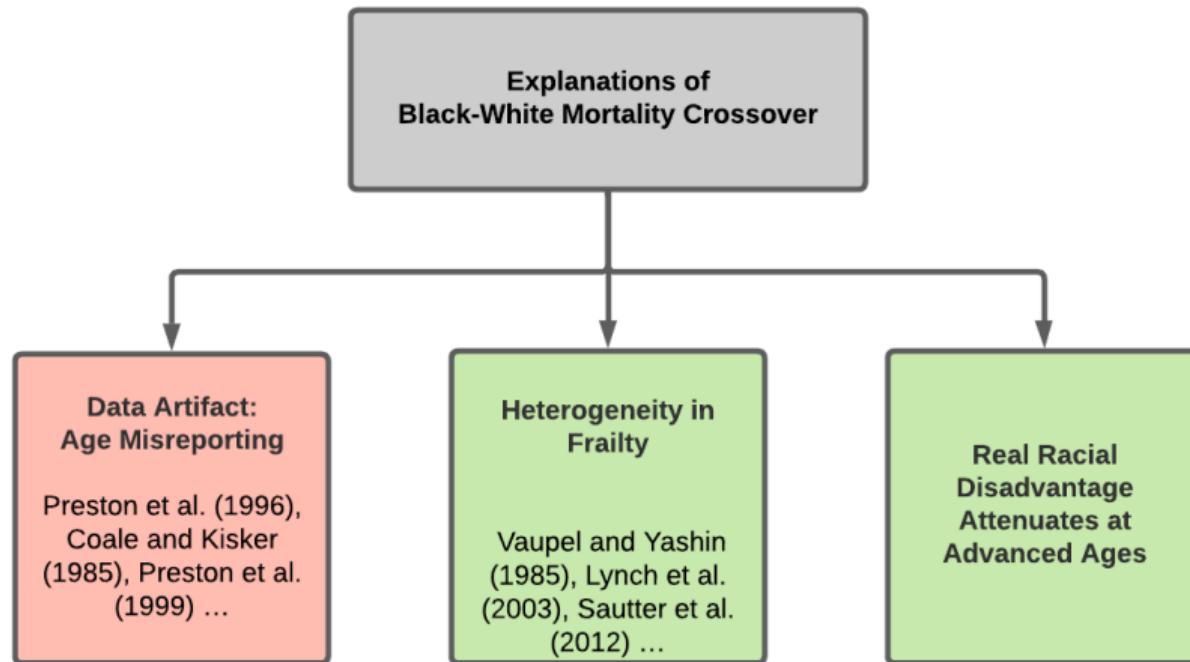
Why understanding Black-White crossover is important

1. Fundamental to our understanding of **inequality** over the life course

2. Oldest-old is fastest growing age segment, potential social policy implications

3. Implications for theories of cumulative disadvantage and weathering (Geronimus, 1992)

Black-White crossover repeatedly documented

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 (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, Neighborhoods		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)

Introduction
○○○●○○○○○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○○

References

Still no consensus on explanation…

**Explanations of
Black-White Mortality Crossover**

**Data Artifact:
Age Misreporting**

Preston et al. (1996),
Coale and Kisker
(1985), Preston et al.
(1999) …

**Heterogeneity in
Frailty**

Vaupel and Yashin
(1985), Lynch et al.
(2003), Sautter et al.
(2012) …

**Real Racial
Disadvantage
Attenuates at
Advanced Ages**

Introduction
○○○○●○○○○○○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○

References

Research questions

1. Is the Black-White mortality crossover a **data artifact**?

2. Does **heterogeneity in frailty** explain the Black-White crossover? Is there really observable late-life mortality selection?

Introduction
ooooo●ooooooooooooooo

Data artifact
ooooo

Frailty
ooooooooo

Empirical testing
ooooooooooooooooo

Discussion
ooooooooooooo

References

New Data Allows Us to Make Progress

▶ Data limitations have hampered efforts to explain crossover

▶ Comparative advantage:
  1. Massive sample (1M deaths)
  2. Cohorts
  3. Covariates

scientific **data**

Check for updates

OPEN

DATA DESCRIPTOR
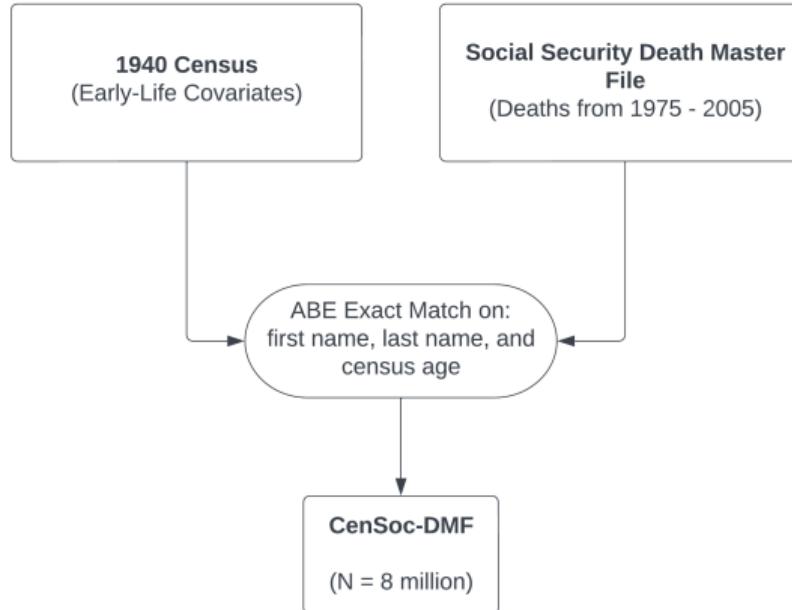
CenSoc: Public Linked Administrative Mortality Records for Individual-level Research

Casey F. Breen[1,2 ✉], Maria Osborne[1] & Joshua R. Goldstein[1 ✉]
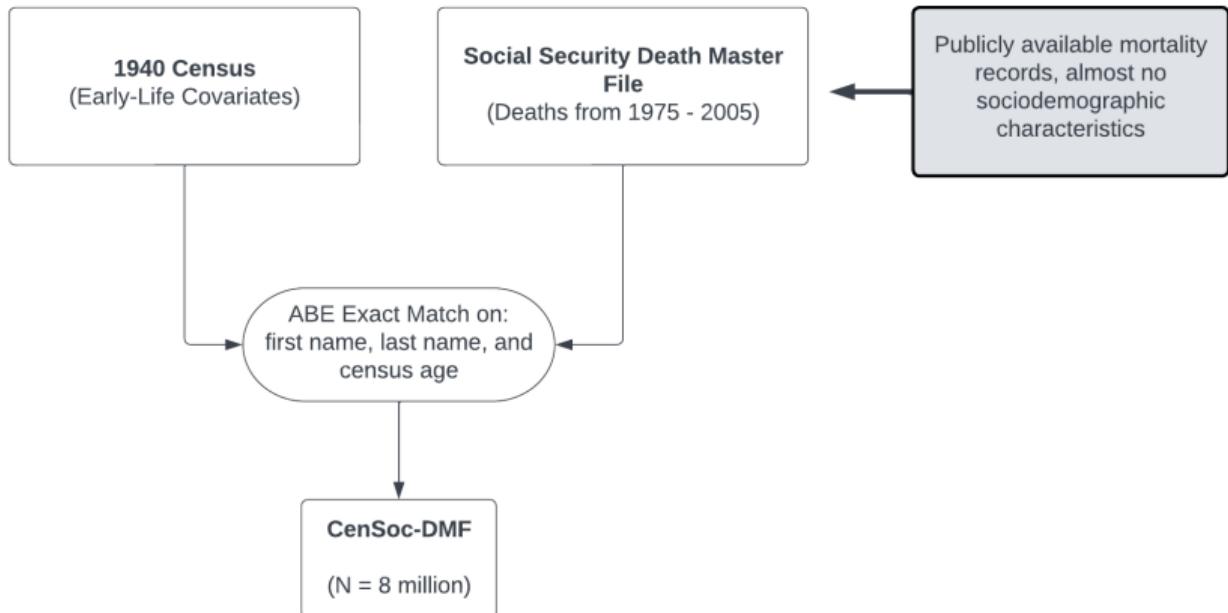
In the United States, much has been learned about the determinants of longevity from survey data and aggregated tabulations. However, the lack of large-scale, individual-level administrative mortality records has proven to be a barrier to further progress. We introduce the CenSoc datasets, which link the complete-count 1940 U.S. Census to Social Security mortality records. These datasets—CenSoc-DMF (N = 4.7 million) and CenSoc-Numident (N = 7.0 million)—primarily cover deaths among individuals aged 65 and older. The size and richness of CenSoc allows investigators to make new discoveries into geographic, racial, and class-based disparities in old-age mortality in the United States. This article gives an overview of the technical steps taken to construct these datasets, validates them using external aggregate mortality data, and discusses best practices for working with these datasets. The CenSoc datasets are publicly available, enabling new avenues of research into the determinants of mortality disparities in the United States.

Breen, Osborne, Goldstein 2023
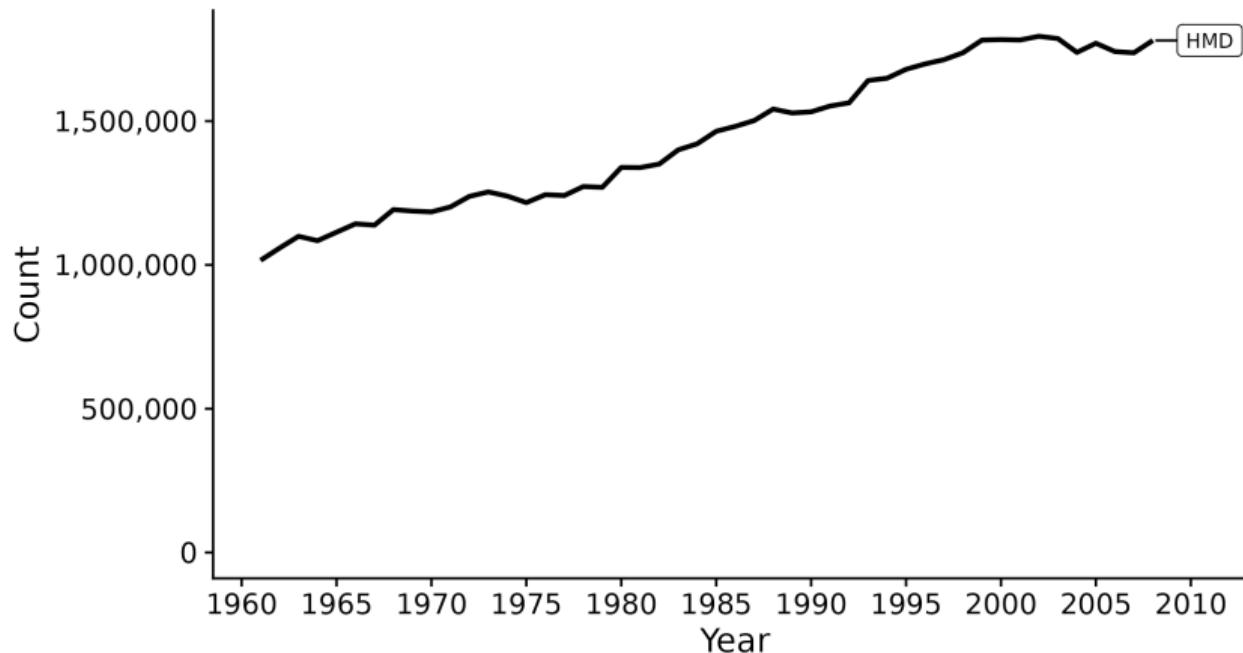
Introduction
○○○○○○●○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○

Empirical testing
○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○○

References

CenSoc-DMF: Linked IPUMS 1940 Census and mortality records

1940 Census
(Early-Life Covariates)

Social Security Death Master File
(Deaths from 1975 - 2005)

ABE Exact Match on:
first name, last name, and census age

CenSoc-DMF

(N = 8 million)

Introduction
○○○○○○○●○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○

References

CenSoc-DMF: Linked IPUMS 1940 Census and mortality records

Introduction
○○○○○○○○○●○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○
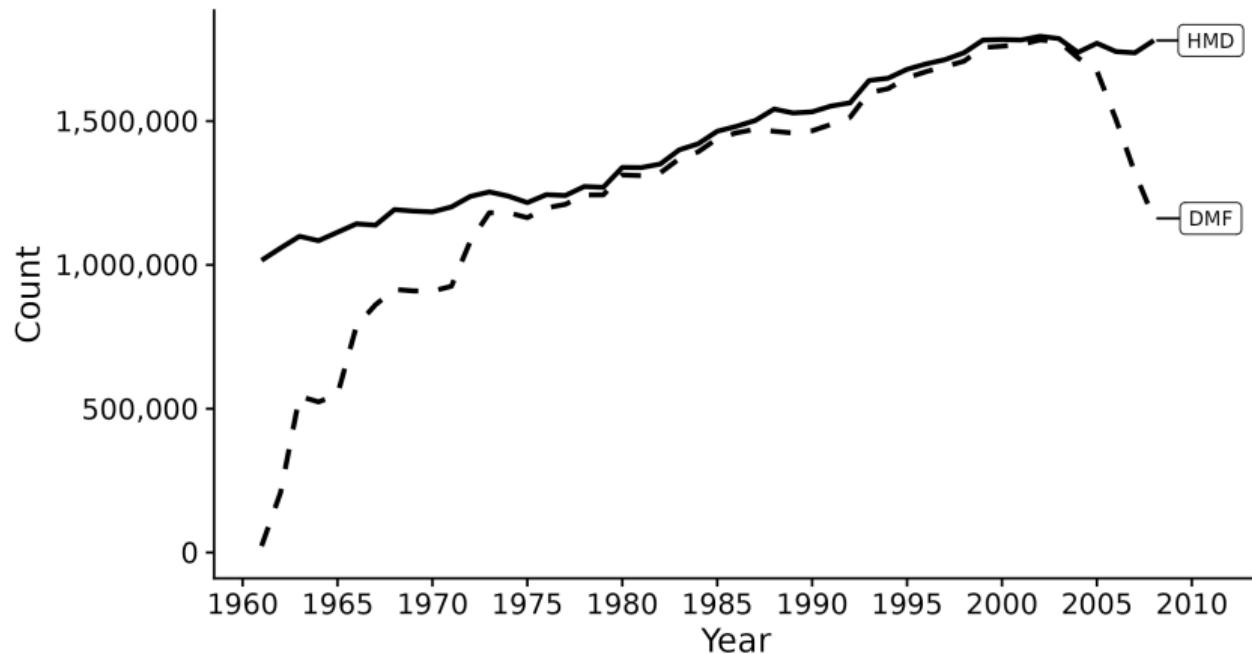
References
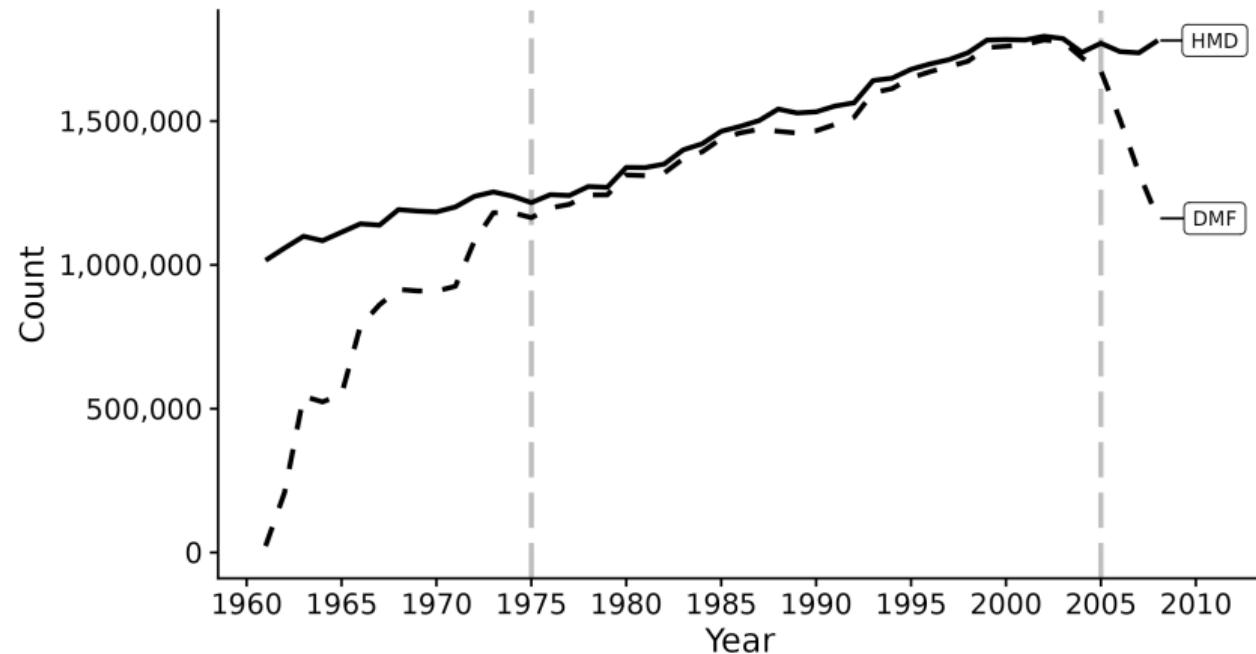
Death Master File (DMF) coverage (65+)
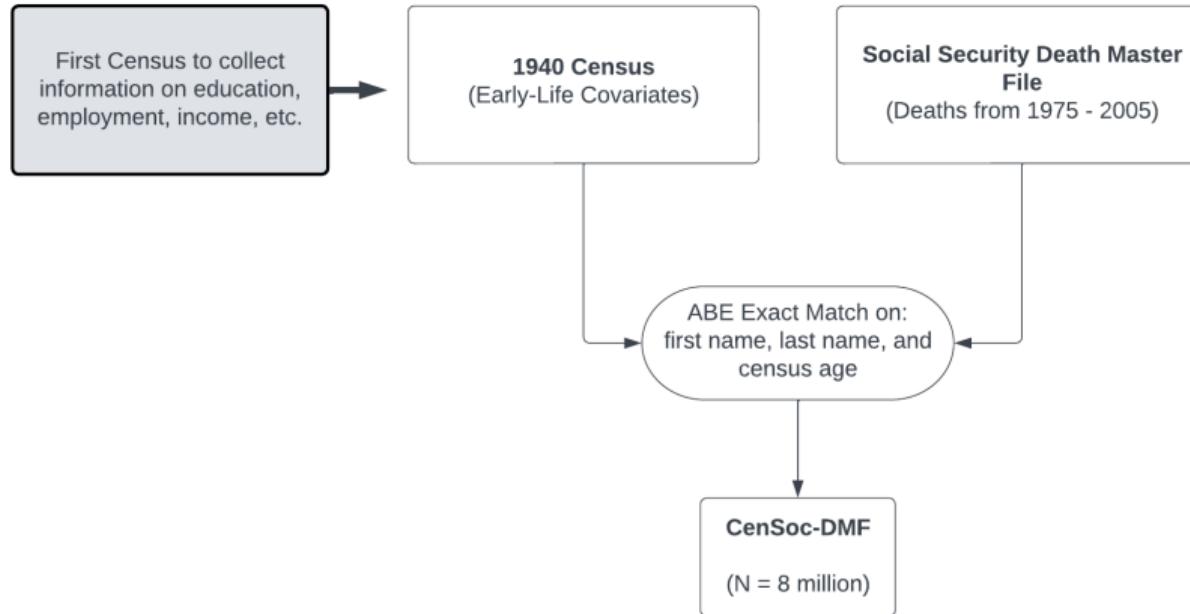
Death Master File (DMF) coverage (65+)

95% death coverage 1975-2005 (65+)

CenSoc-DMF: Linked 1940 Census and mortality records

Introduction
○○○○○○○○○○○○○●○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○

References

1940 Census

▶ 1940 Census reflected heightened time of social awareness brought about by Great Depression

1940 Census

▶ 1940 Census reflected heightened time of social awareness brought about by Great Depression

▶ First decennial census to include question on educational attainment, wage and salary income, and detailed questions on employment

Introduction
○○○○○○○○○○○○○○●○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○

References
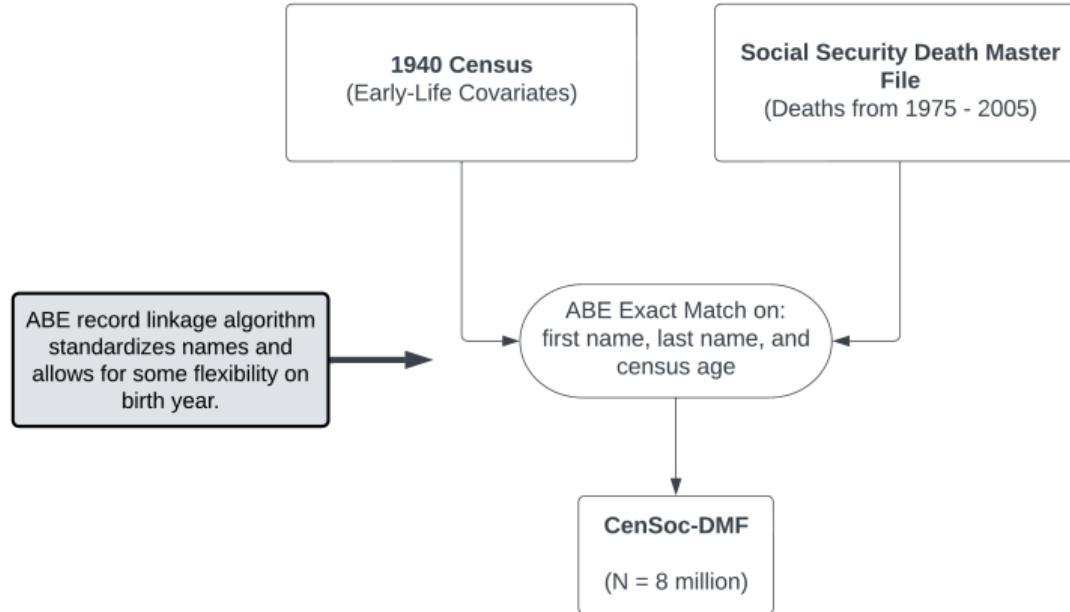
1940 Census

▶ 1940 Census reflected heightened time of social awareness brought about by Great Depression

▶ First decennial census to include question on educational attainment, wage and salary income, and detailed questions on employment
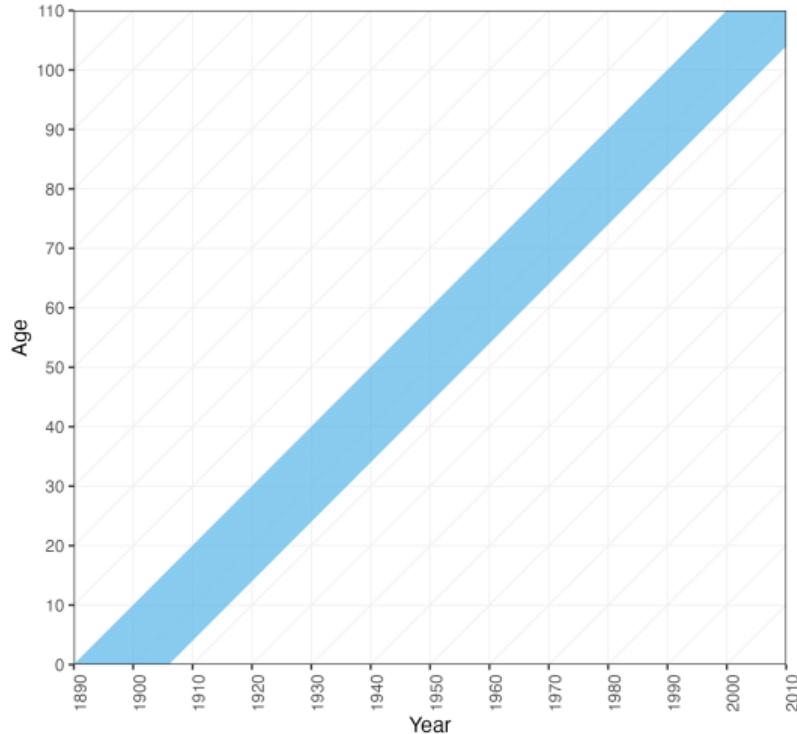
1940 Census Form

CenSoc-DMF: Linked census and mortality records

1940 Census
(Early-Life Covariates)

Social Security Death Master File
(Deaths from 1975 - 2005)

ABE record linkage algorithm standardizes names and allows for some flexibility on birth year.

ABE Exact Match on:
first name, last name, and census age

CenSoc-DMF

(N = 8 million)

Introduction
○○○○○○○○○○○○○○○●○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○○

References

Analytic Sample

- Birth cohorts of 1890-1905
    - $N = 900,000$ deaths

- Sample restrictions
    - Men only

    - U.S. born

Introduction
○○○○○○○○○○○○○○○●○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○

References
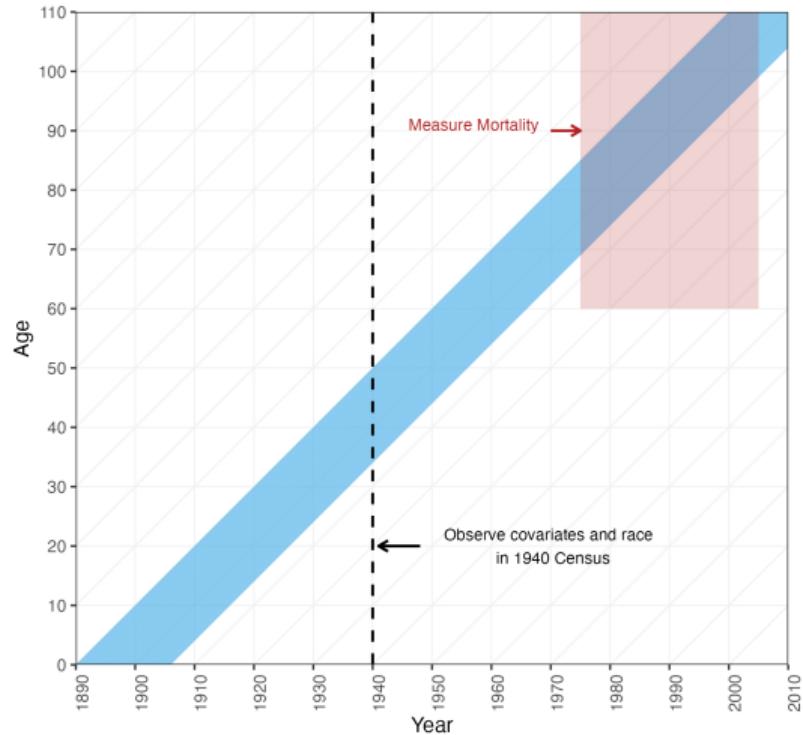
Analytic Sample

- Birth cohorts of 1890-1905
  - $N = 900,000$ deaths

- Sample restrictions
  - Men only

  - U.S. born

Analytic Sample

- Birth cohorts of 1890-1905
  - $N = 900,000$ deaths

- Sample restrictions
  - Men only

  - U.S. born

Introduction
ooooooooooooooooooo●oo

Data artifact
ooooo

Frailty
ooooooooo

Empirical testing
ooooooooooooooo

Discussion
ooooooooooo

References

Analytic Sample

▶ Birth cohorts of 1890-1905
  ▶ $N = 900,000$ deaths

▶ Sample restrictions
  ▶ Men only
  ▶ U.S. born

Birth cohorts of 1890-1905: extinct cohort method
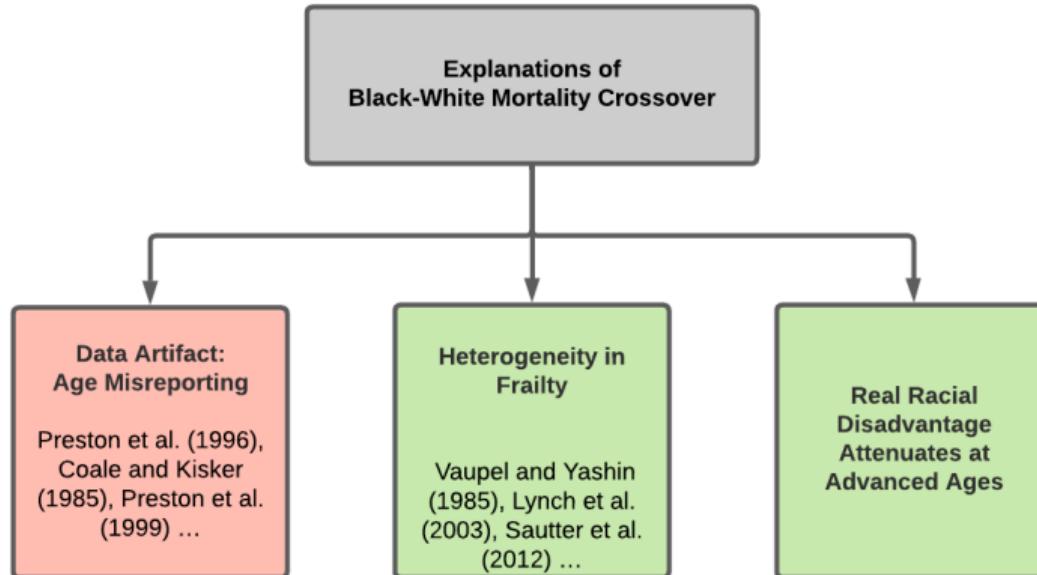
▶ Assumes that all members of the cohort have died by a certain year

▶ Uses recorded deaths over time to reconstruct the cohort's survival pattern

$$q_x = \frac{d_x}{\sum_x^\infty d_i} = \frac{d_x}{l_x} \tag{1}$$
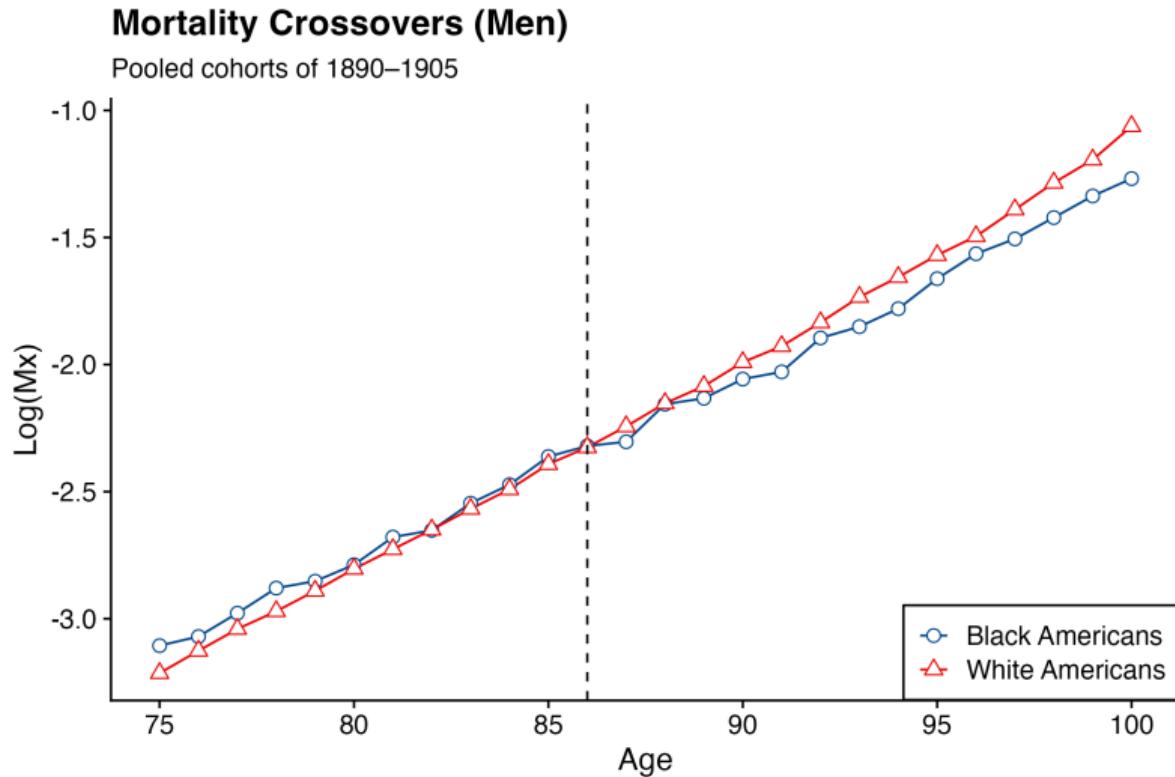
where:

▶ q_x is the probability of dying at age x.

▶ d_x is number of deaths at age x

Question 1: Is the crossover a data artifact?

Introduction
○○○○○○○○○○○○○○○○○○○○○○

Data artifact
●○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○○

References

Black-White crossover (extinct cohort method)

Mortality Crossovers (Men)

Pooled cohorts of 1890–1905

Introduction

Data artifact

Frailty

Empirical testing

Discussion

References
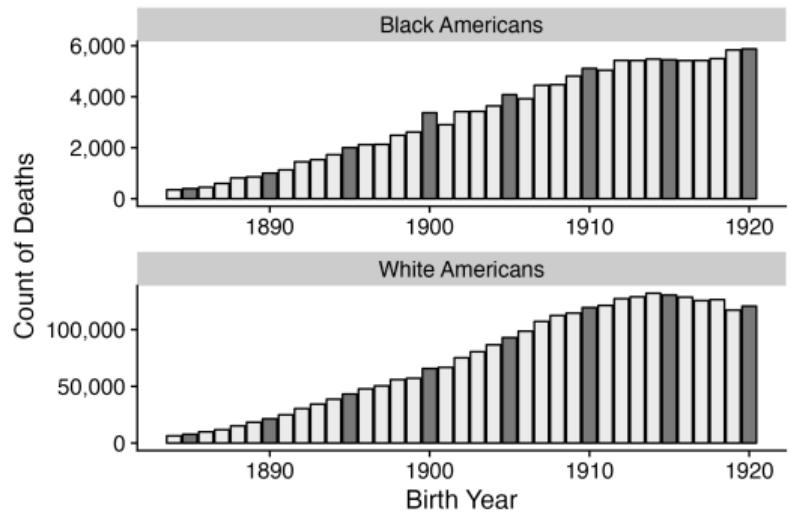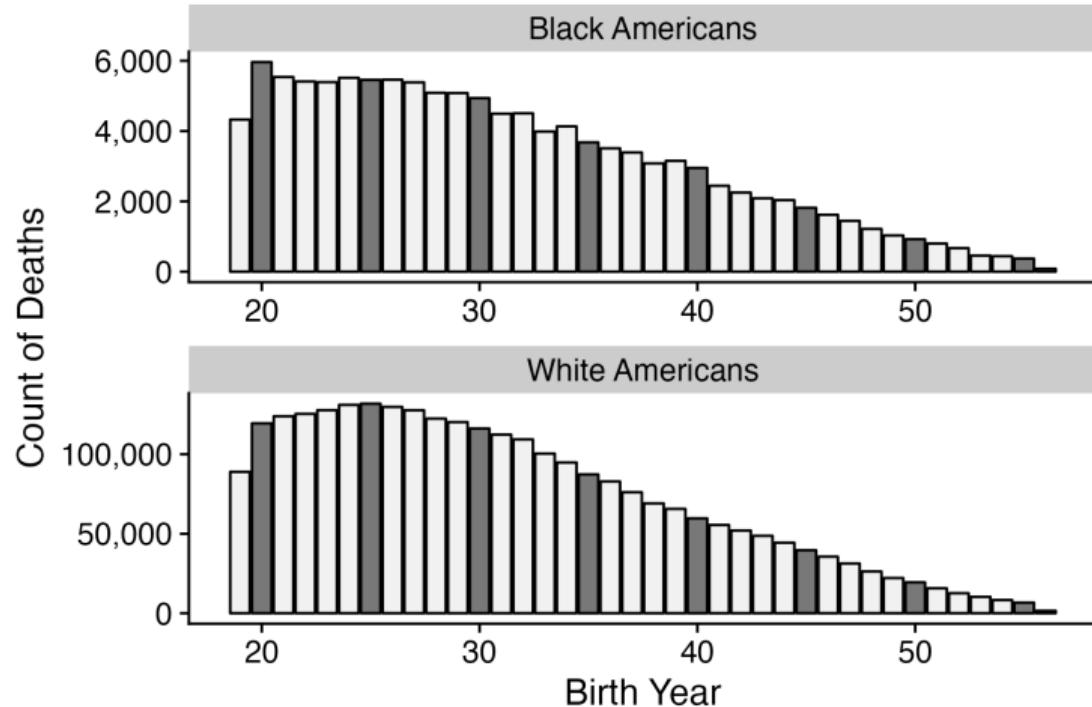
Question 1: Is the crossover a data artifact?

Background: Age of death calculated from date of birth and date of death
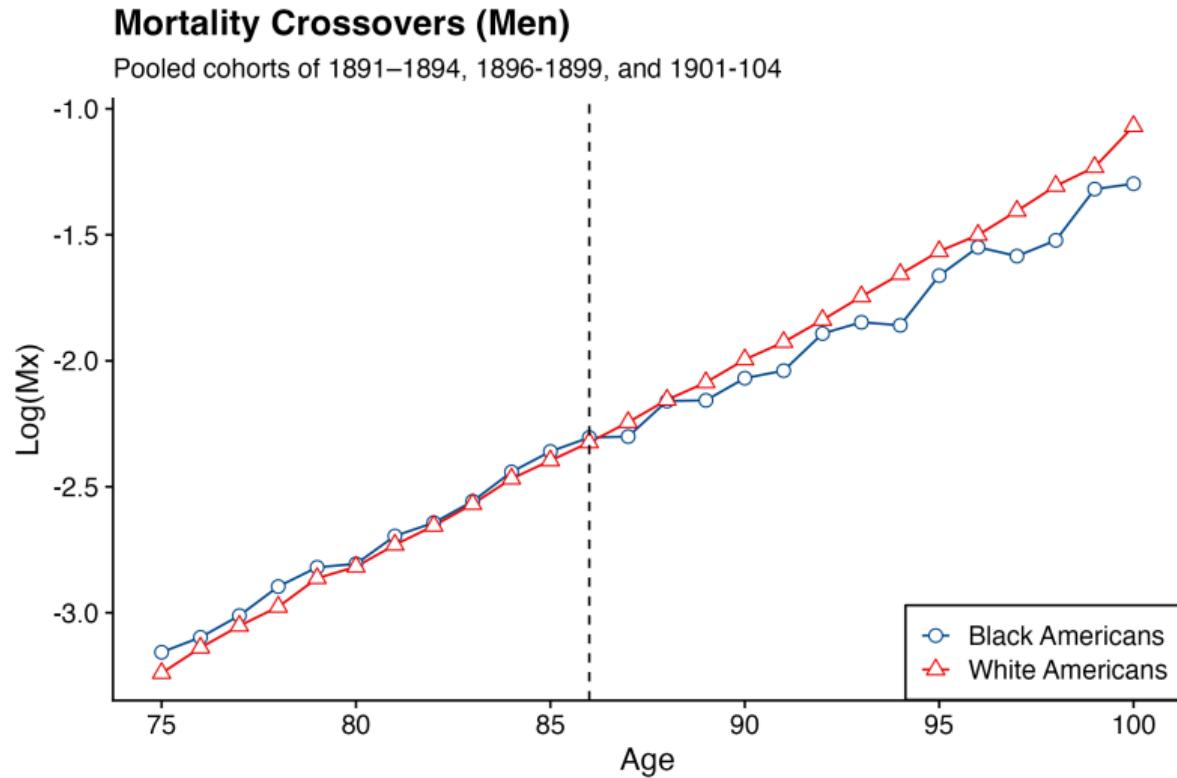
1. Minimal age heaping on birth year...

2. Date of death gets reported immediately (no heaping)

3. Institutional incentive: Social Security wants to accurately track birth date

4. Linkage requires close match on year of birth and Census age
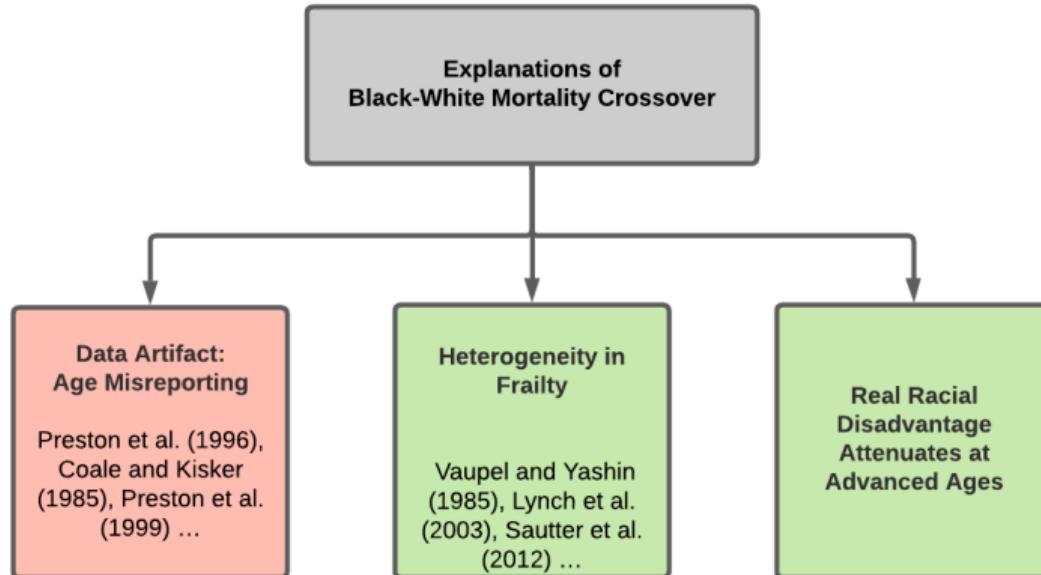
No age heaping in 1940 Census

Black-white crossover (extinct cohort method)

Mortality Crossovers (Men)

Pooled cohorts of 1891–1894, 1896-1899, and 1901-104

Legend:
- Black Americans
- White Americans

Question 2: Is the crossover driven by heterogeneity in frailty?

Heterogeneity in frailty – overview of approach

▶ First, how much **mortality selection** do we actually observe…?

▶ Second, does stratifying on heterogeneity uncross the crossover?

HETEROGENEITY'S RUSES: SOME SURPRISING EFFECTS OF SELECTION ON POPULATION DYNAMICS

James W. Vaupel and Anatoli I. Yashin
Population Program, International Institute for Applied Systems Analysis, Laxenburg, Austria

Unpacking the black box of frailty...

▶ **Frailty:** an individual's susceptibility to death
  ▶ Wealth, education, environmental, behavioral, etc.

▶ Lots of theorizing on frailty — but less empirical evidence due to data limitations

▶ Data-driven investigation of components of frailty that we can observe: sociodemographic characteristics
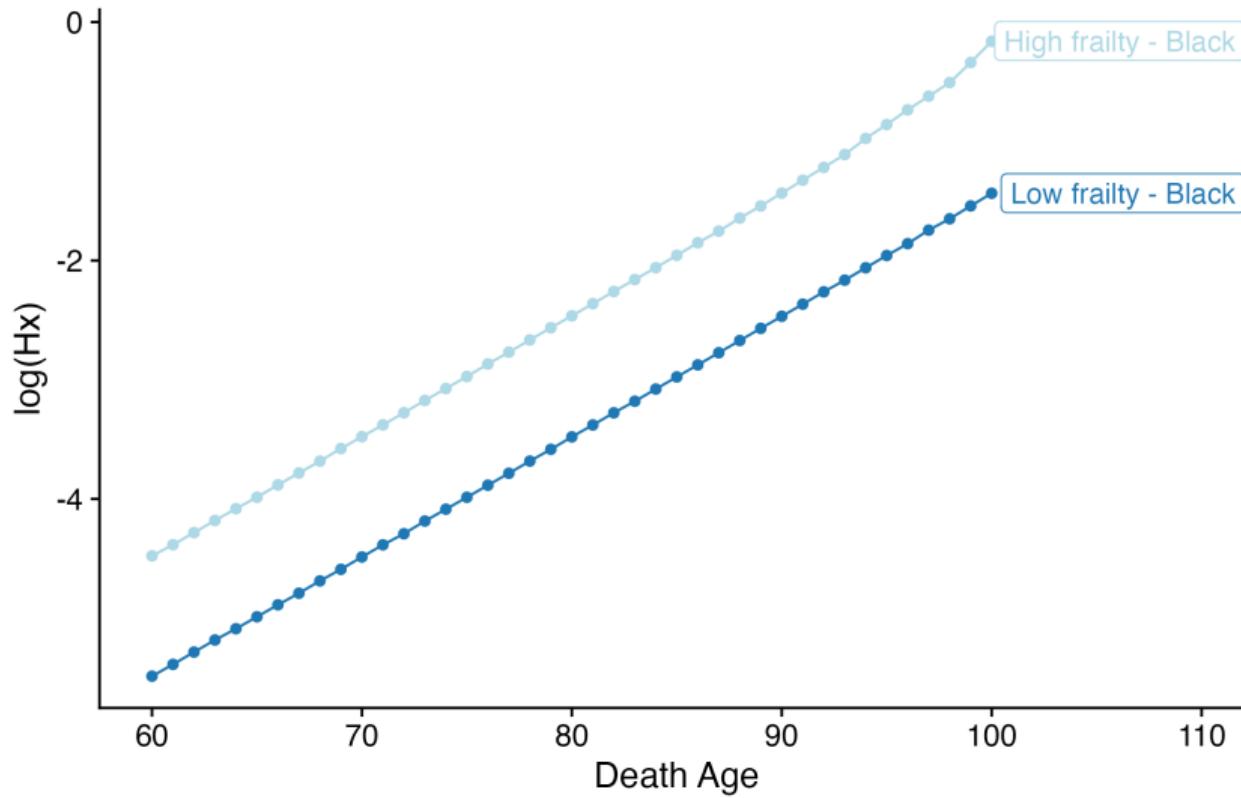
The theory of heterogeneity in frailty

▶ To get a crossover, higher initial mortality population must have **higher variance** in frailty
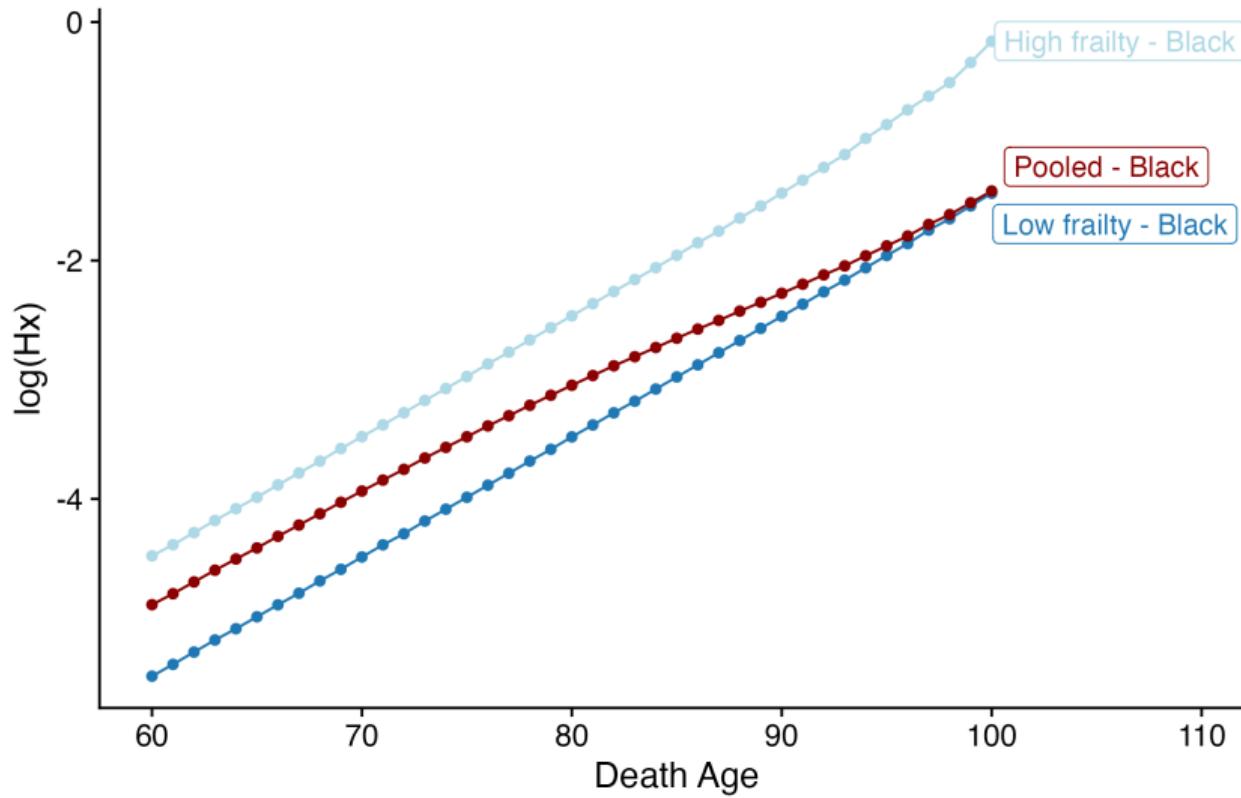
The theory of heterogeneity in frailty

▶ To get a crossover, higher initial mortality population must have **higher variance** in frailty

▶ As the cohorts age, mortality selection is much stronger for the high mortality, high variance group

▶ So much stronger, that eventually the frailty of survivors actually crosses
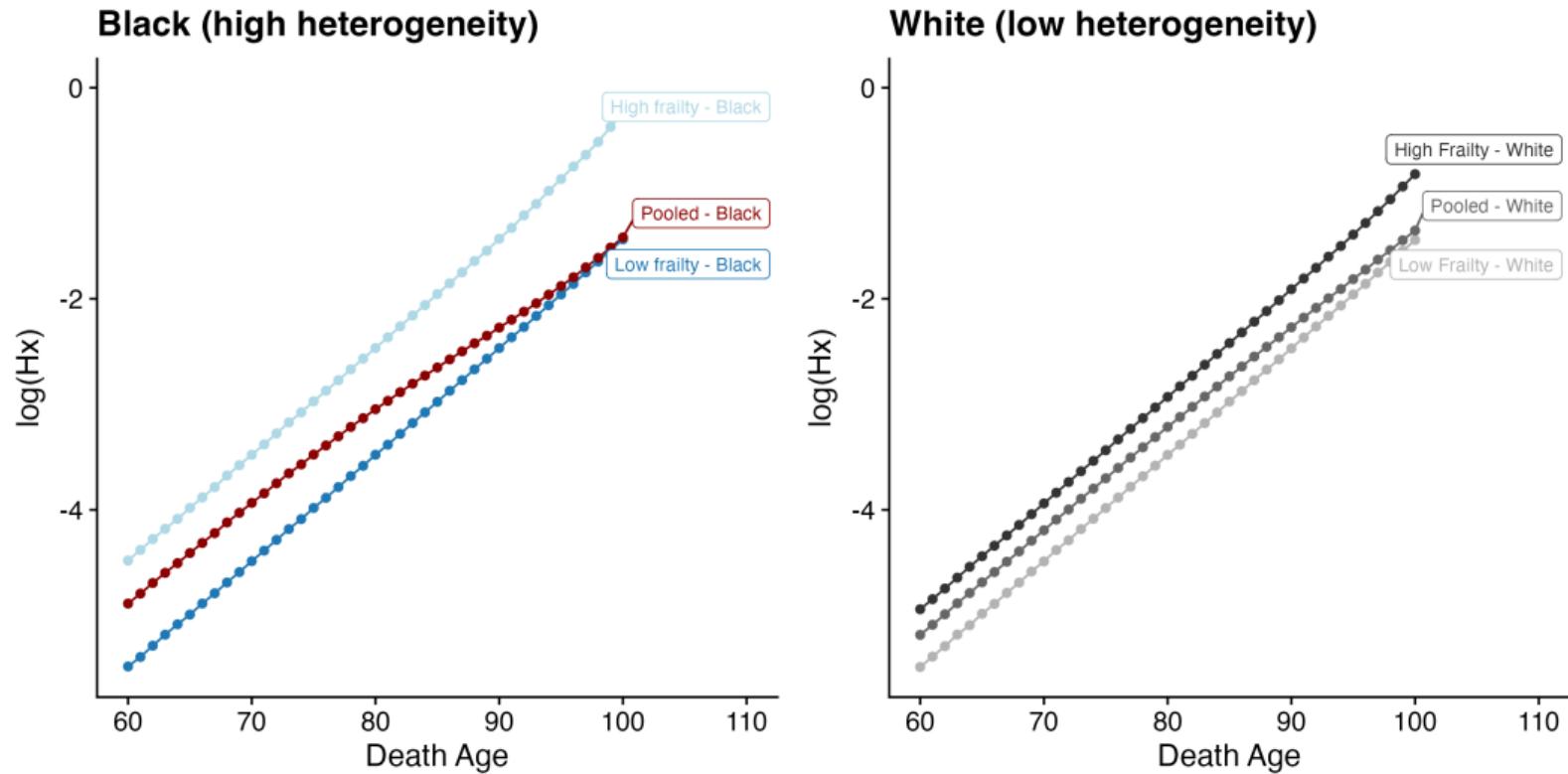
Heterogeneity in frailty (stylized example)
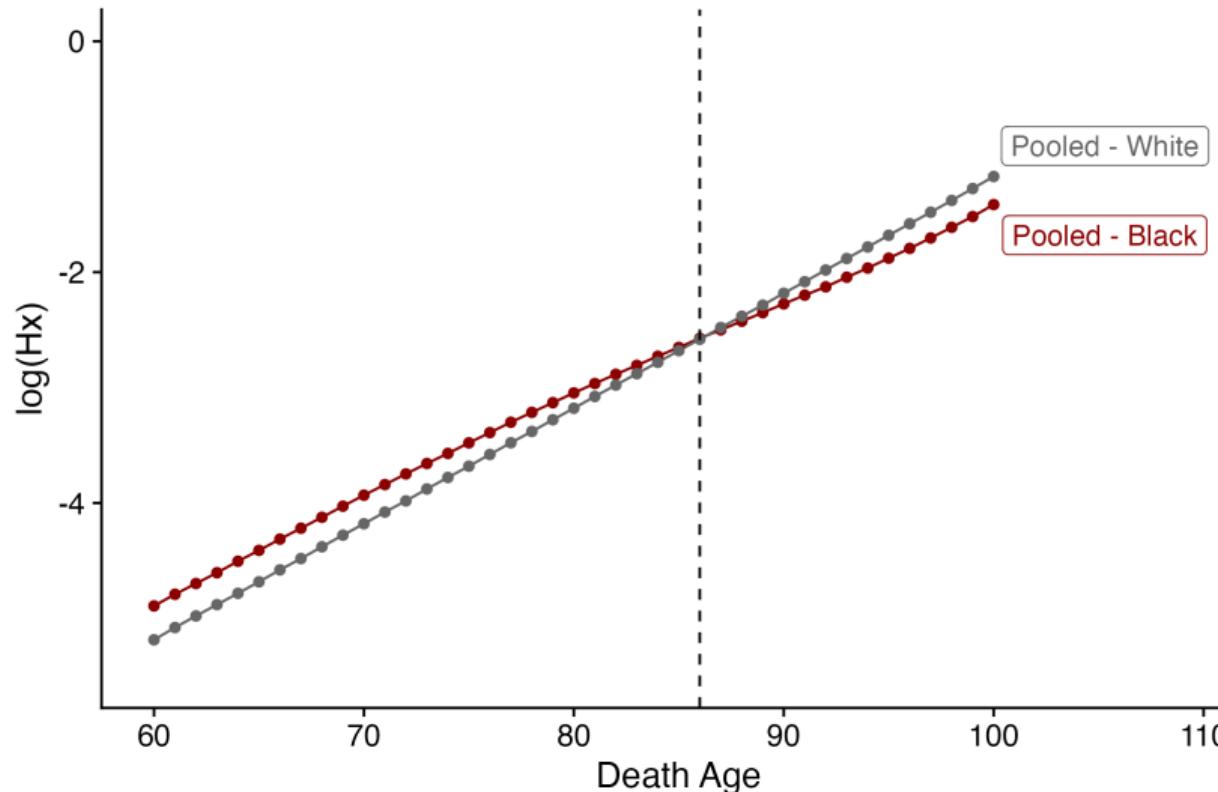
Heterogeneity in frailty (stylized example)

Heterogeneity in frailty (stylized example)

Black (high heterogeneity)

White (low heterogeneity)

Introduction
Data artifact
Frailty
Empirical testing
Discussion
References
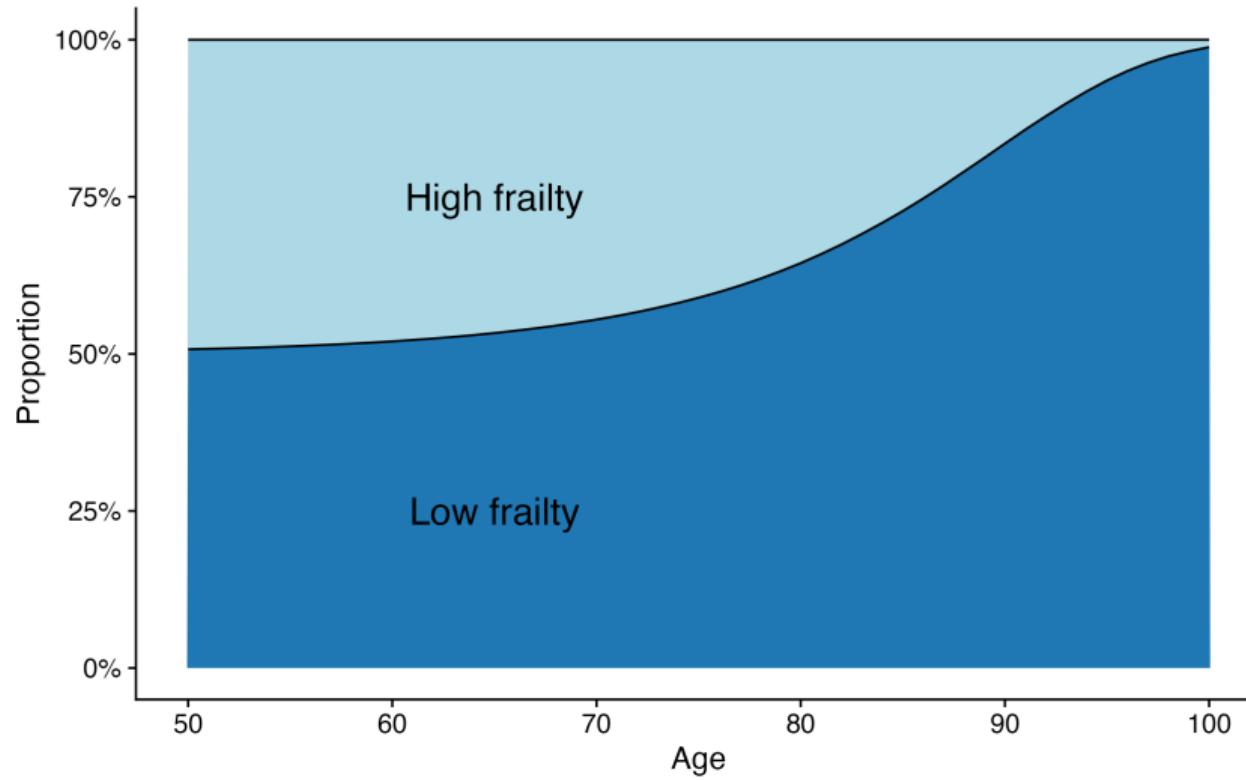
Heterogeneity in frailty (stylized example)

Very dramatic shift in composition of survivors...

Empirical testing

▶ In order for the theory of frailty to explain the Black-White crossover, there must be stronger frailty selection over the life course for Black Americans than for White Americans

Empirical testing

▶ In order for the theory of frailty to explain the Black-White crossover, there must be stronger frailty selection over the life course for Black Americans than for White Americans

▶ We don't observe frailty but we observe characteristics that comprise part of frailty

    ▶ Educational attainment

    ▶ Occupation and income

   ▶ Wealth

Changing educational composition of survivors

# Changing composition of survivors

Educational Attainment
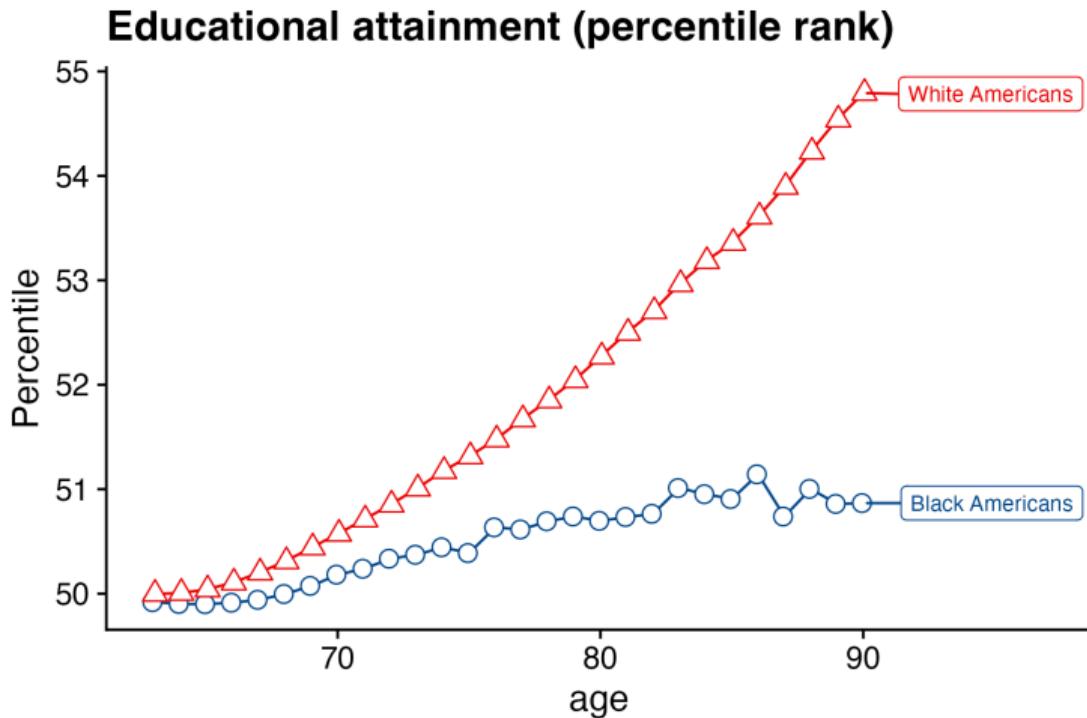
Changing educational composition of survivors (percentile)

**Educational attainment (percentile rank)**

White Americans

Black Americans

Introduction
○○○○○○○○○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○●○○○○○○○○○○○○

Discussion
○○○○○○○○○○○○
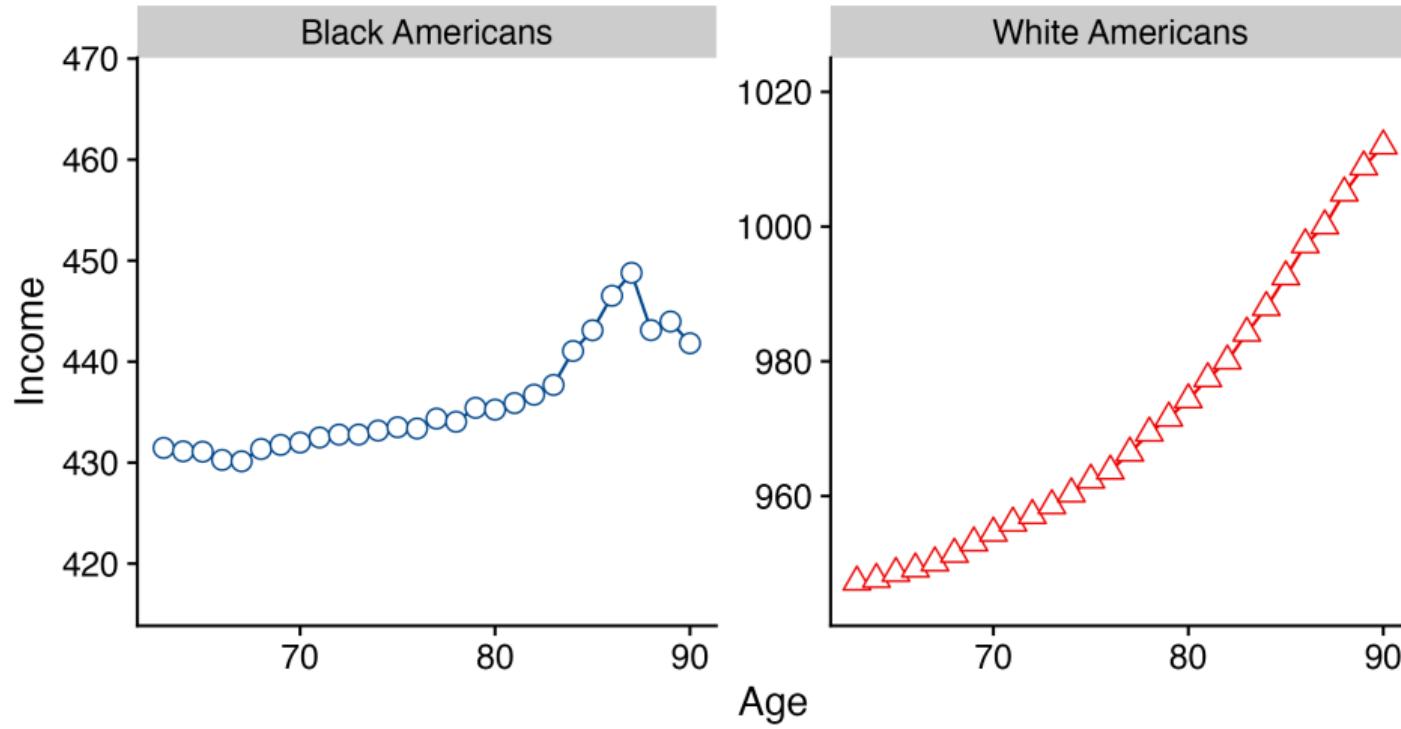
References
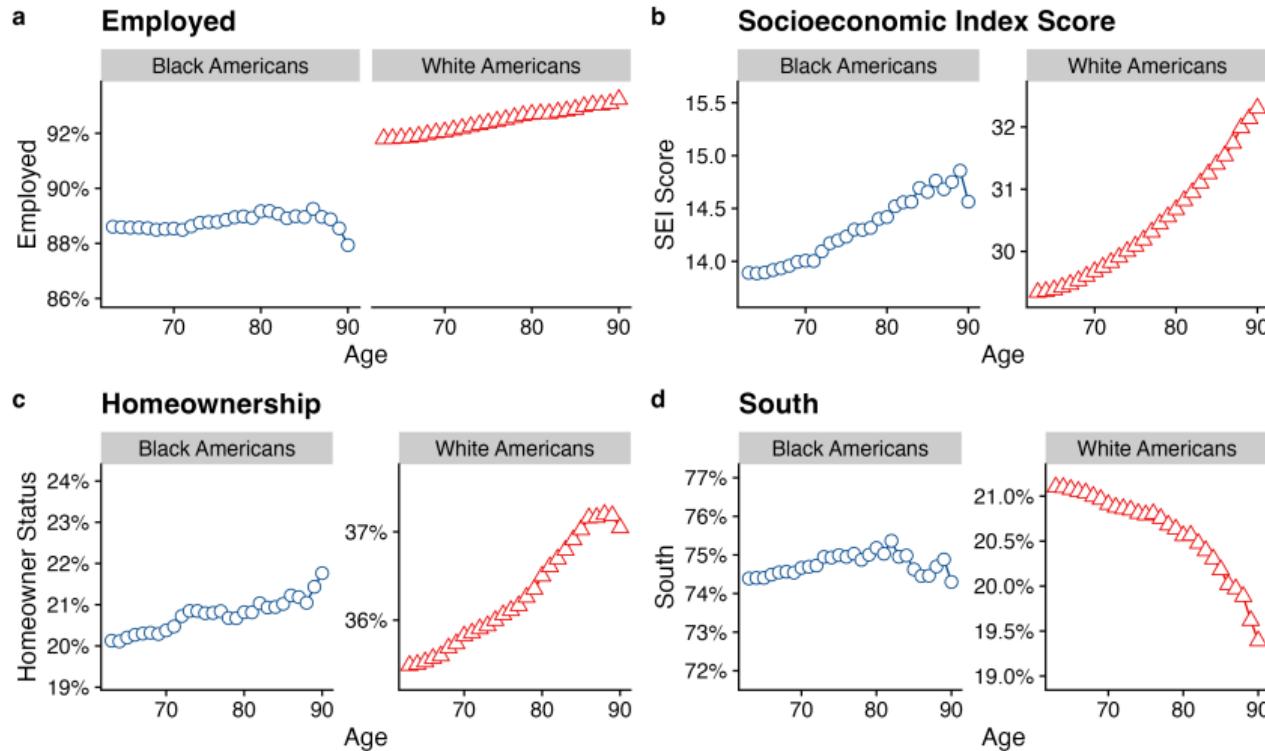
Changing composition of survivors
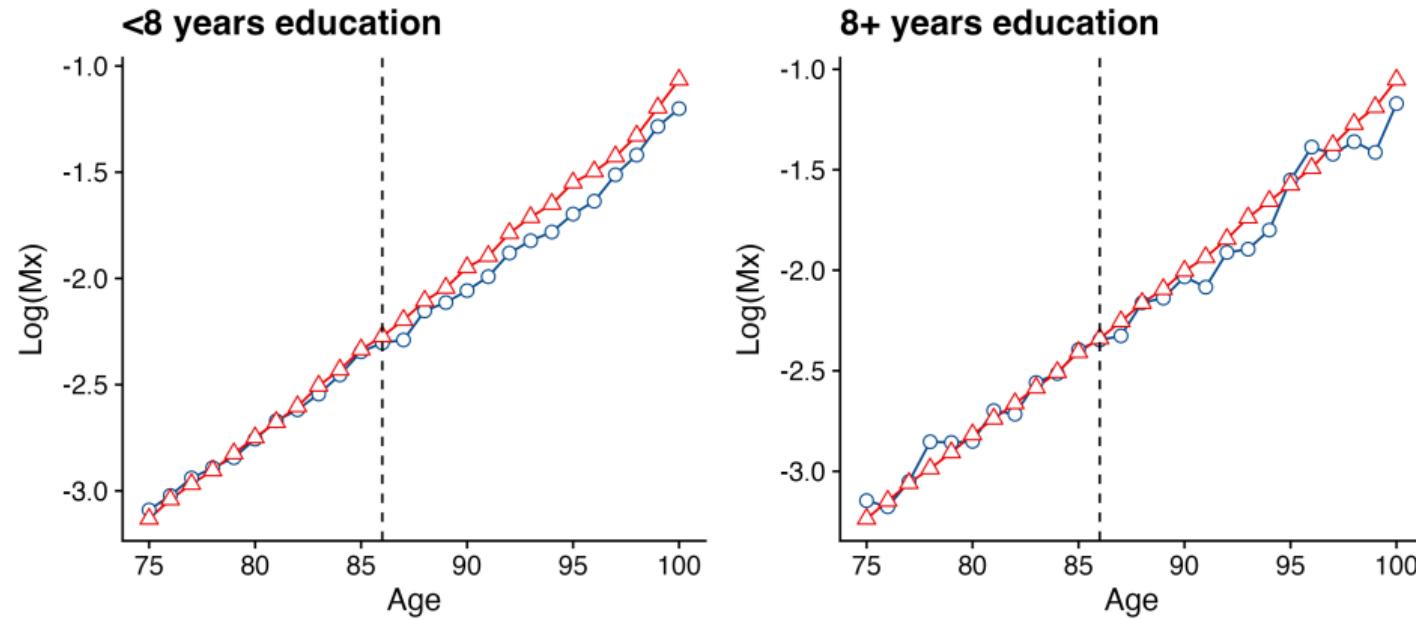
Wage and Salary Income

Changing composition of survivors

a **Employed**

b **Socioeconomic Index Score**

c **Homeownership**

d **South**
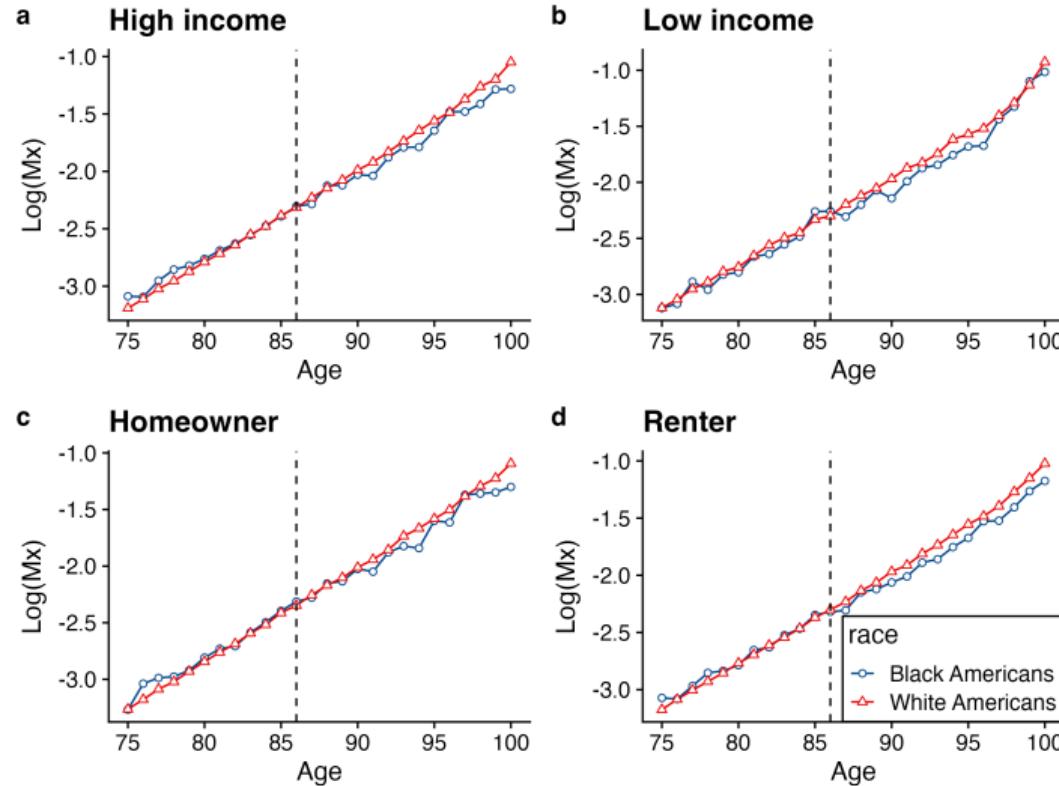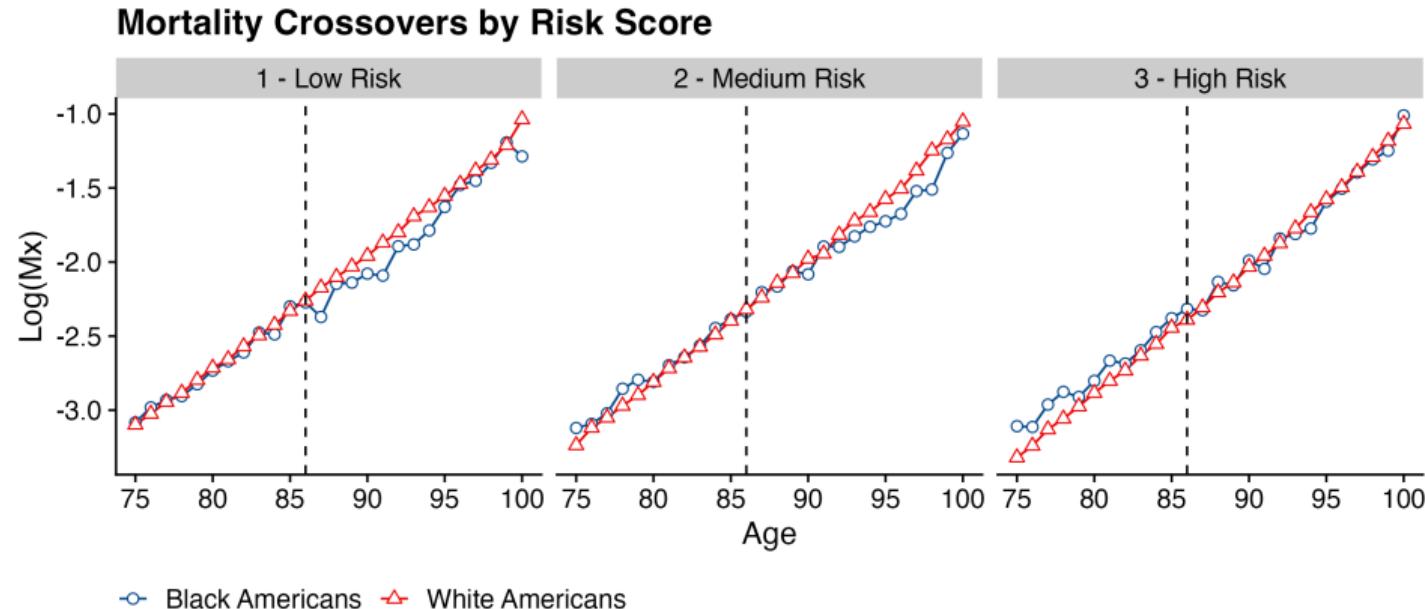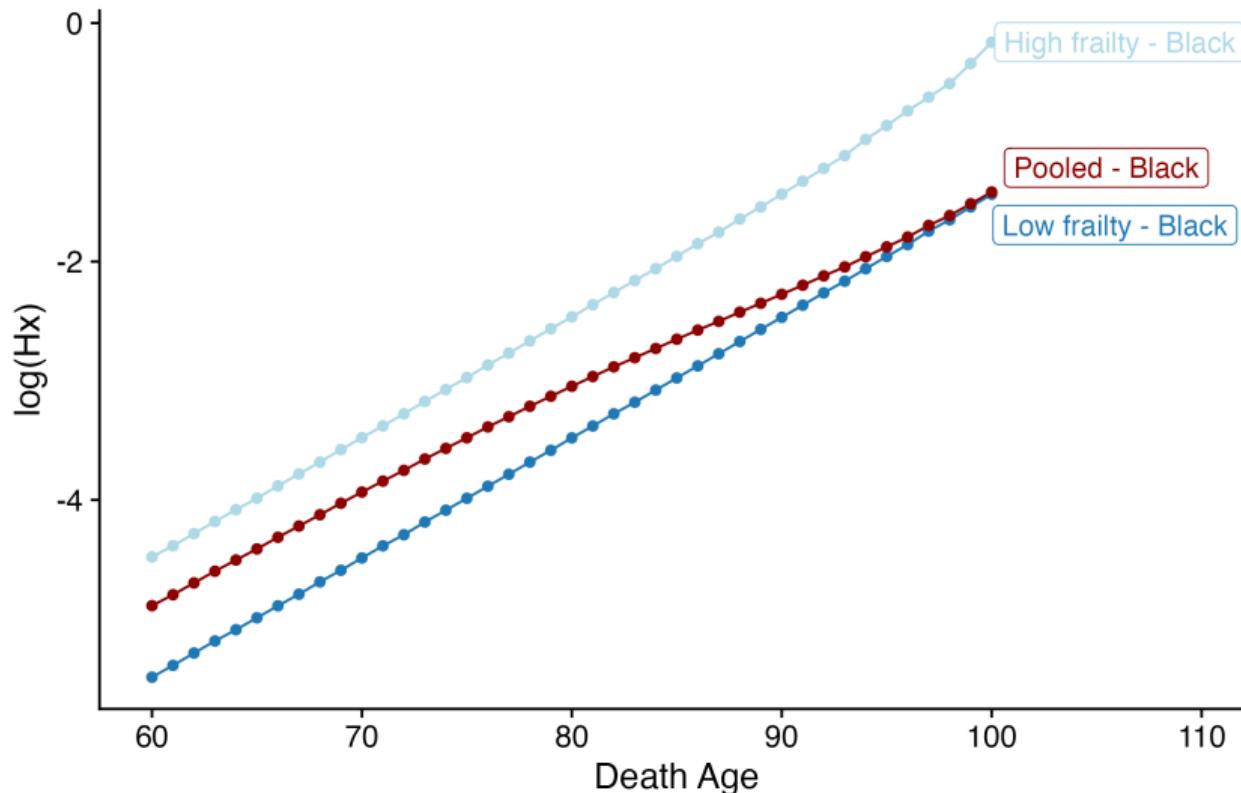
Stratifying by dimensions of frailty

Stratifying by dimensions of frailty

Introduction
○○○○○○○○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○●○○○○○

Discussion
○○○○○○○○○○○○

References

Stratifying on risk score
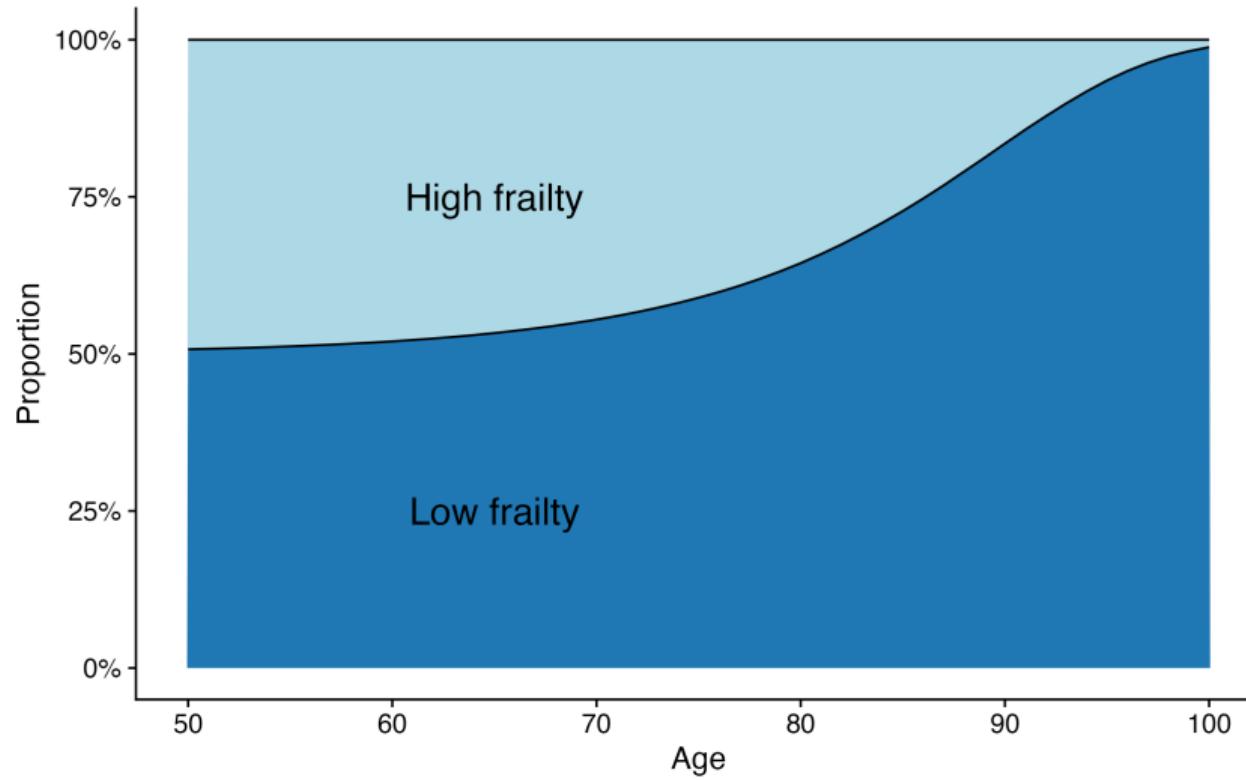
Mortality Crossovers by Risk Score
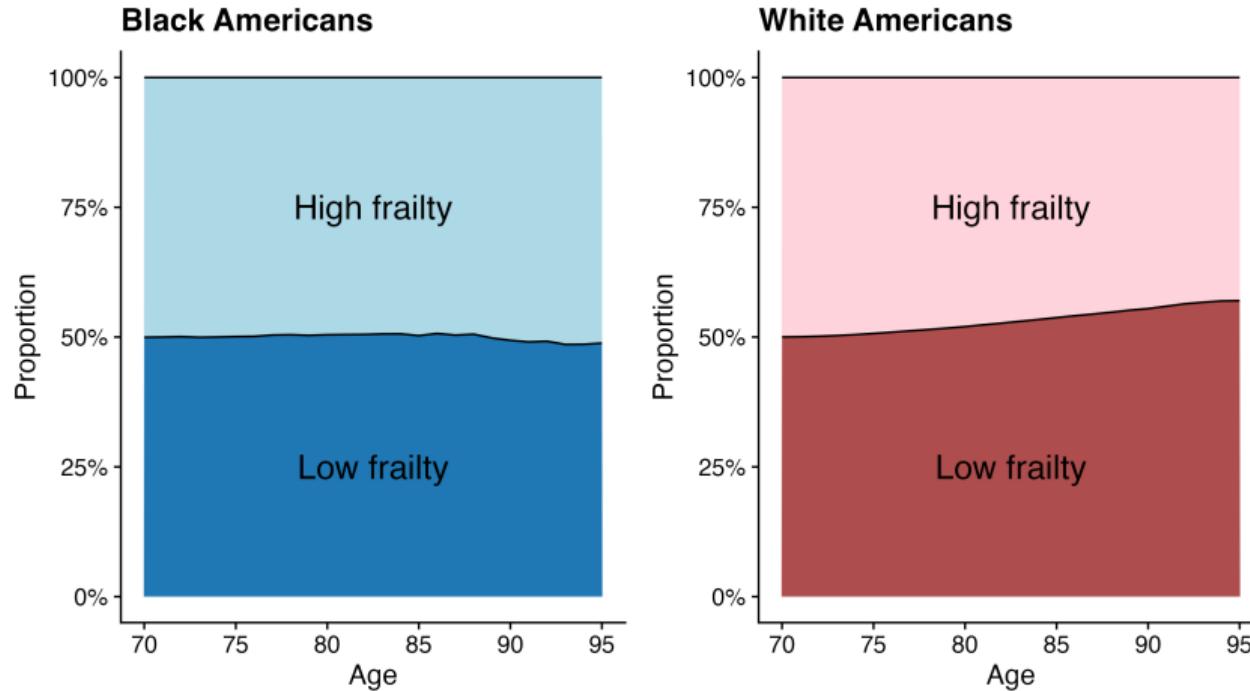
How much selection would we need? (Simulation)
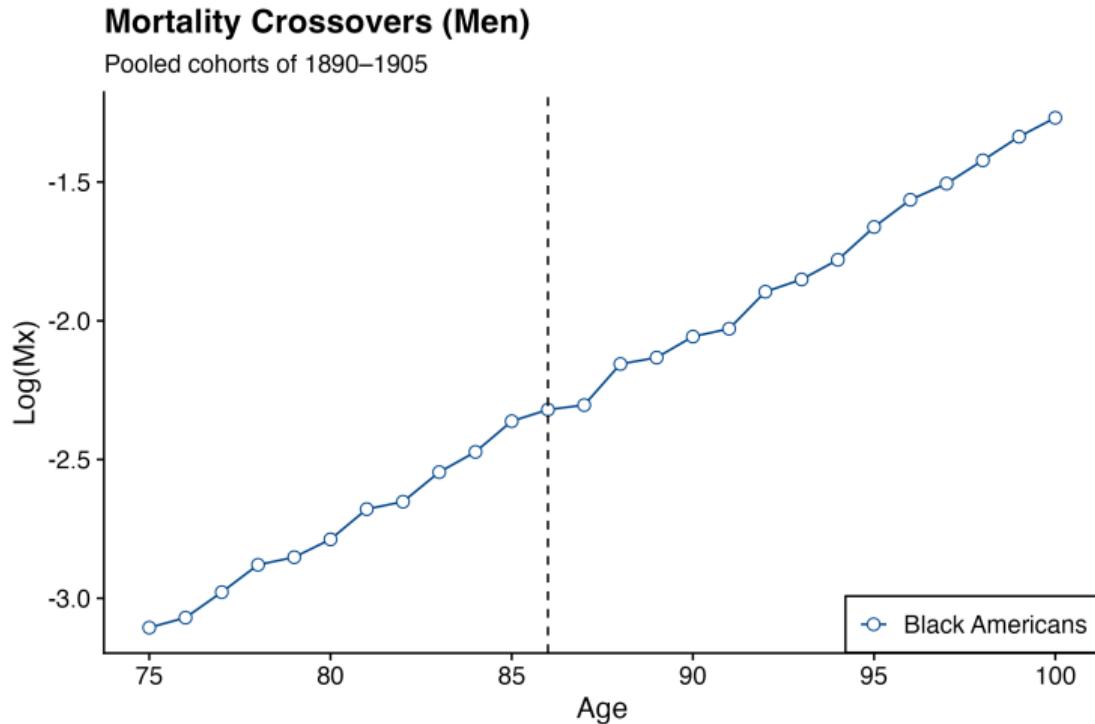
How much selection would we need for Blacks? (Simulation)
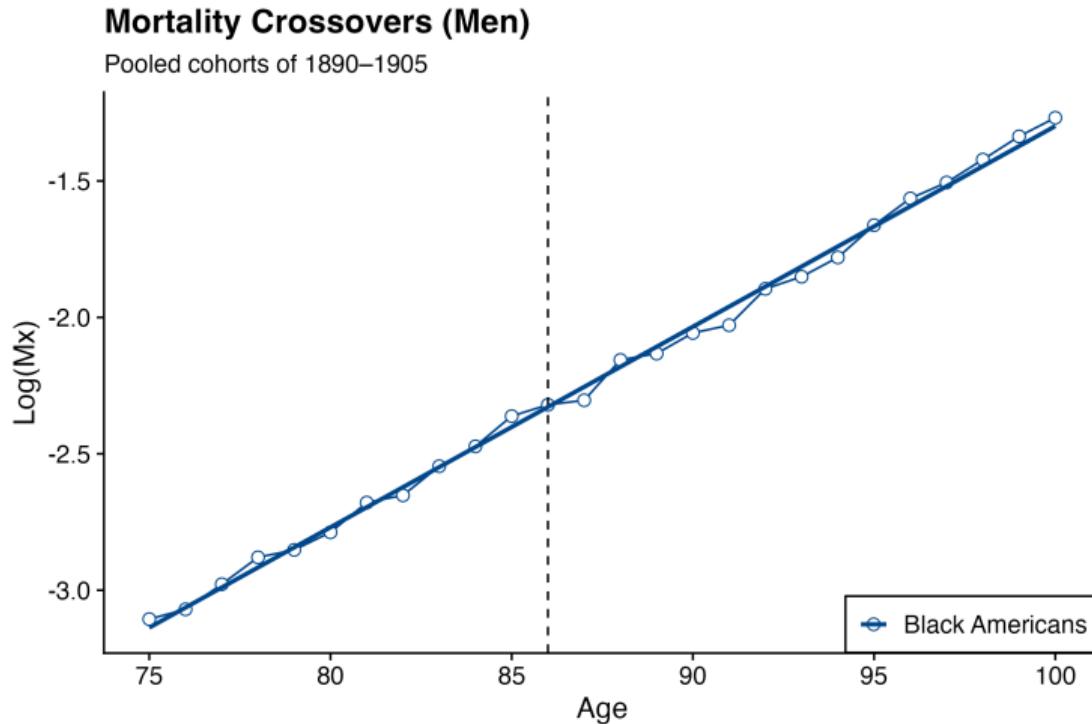
How much selection do we actually observe? Very little

Black Americans / **White Americans** — proportion of High frailty / Low frailty by Age (70–95)

Simple test - visible deceleration? (Bend in curve?)

Mortality Crossovers (Men)
Pooled cohorts of 1890–1905

Perfectly linear…

Mortality Crossovers (Men)
Pooled cohorts of 1890–1905

No support for heterogeneity in frailty explanation
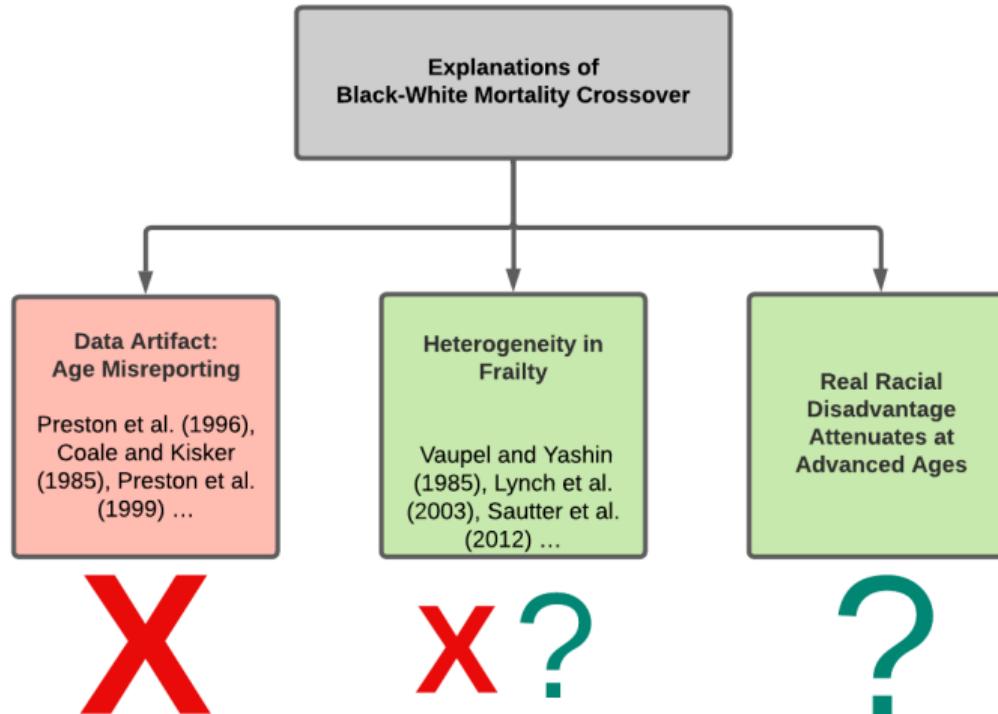
1. Very little mortality selection

No support for heterogeneity in frailty explanation

1. Very little mortality selection

2. Mortality selection stronger for White Americans than Black Americans
   ▶ Cause a widening — not convergence/crossover — of mortality rates

3. No observable mortality deceleration

Revisiting explanations...

Introduction ○○○○○○○○○○○○○○○○○○○○○○○○

Data artifact ○○○○○

Frailty ○○○○○○○○○

Empirical testing ○○○○○○○○○○○○○○○

Discussion ●○○○○○○○○○○○○

References

Where to next? Adaptive resilience

- **Resilience**

  - Lifelong exposure to adversity (racism, economic hardship, discrimination) may cultivate adaptive coping mechanisms

  - Black Americans have higher levels of self-esteem and religiosity than white Americans (**Louie 2024**), better mental health (**Erving 2021**) and subjective survival expectations (**Bernstein and Sasson 2021**)

- **Supportive Social Structures**

  - Older Black adults are more deeply embedded in extended family systems + live closer to kin (**Springer, Crowder, et al. 2023**)

  - More embedded networks provide caregiving, daily help, companionship, protection from loneliness (**Cornwell and Qu, 2024, Umberson et al. 2010, Durkheim 1897**)

Conclusions

▶ Black-White Crossover is real — *not **data artifact***

Conclusions

▶ Black-White Crossover is real — *not **data artifact***

▶ No support for **heterogeneity in frailty** explanation
  ▶ Not educational attainment, not income, not wealth, not risk score
  ▶ However, other dimensions of heterogeneity might be responsible (biomarkers, self-rated health)

Conclusions
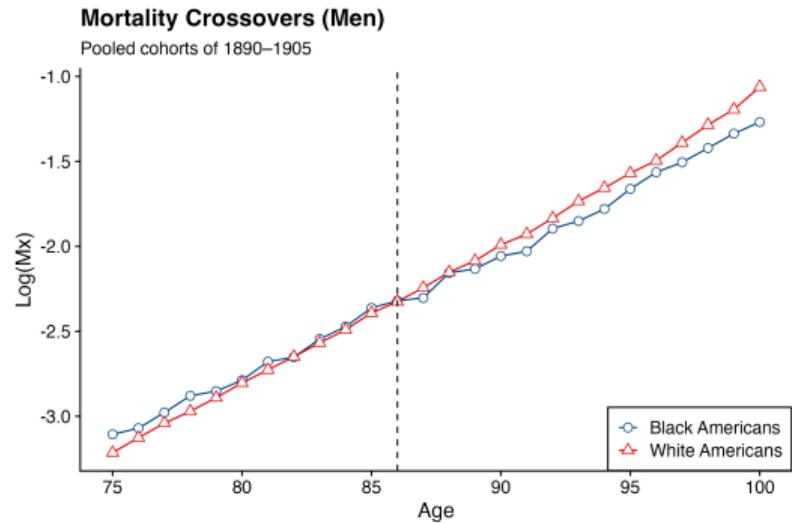
▶ Black-White Crossover is real — *not **data artifact***

▶ No support for **heterogeneity in frailty** explanation
  ▶ Not educational attainment, not income, not wealth, not risk score
  ▶ However, other dimensions of heterogeneity might be responsible
    (biomarkers, self-rated health)

▶ Attenuation of disadvantage from adaptive resilience: **Intriguing**, but need
  more theoretical and empirical work

▶ New explanations are needed (!)

Thank you — questions?
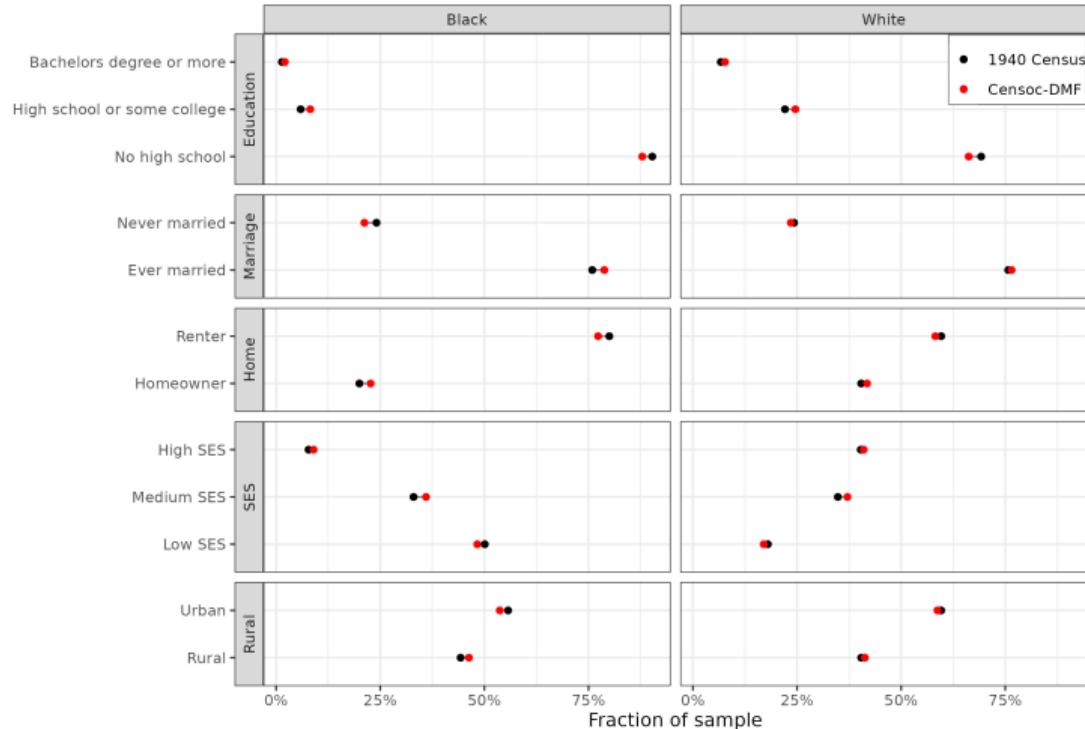
Data: CenSoc.Berkeley.edu

Funding: R01AG058940, R01AG076830
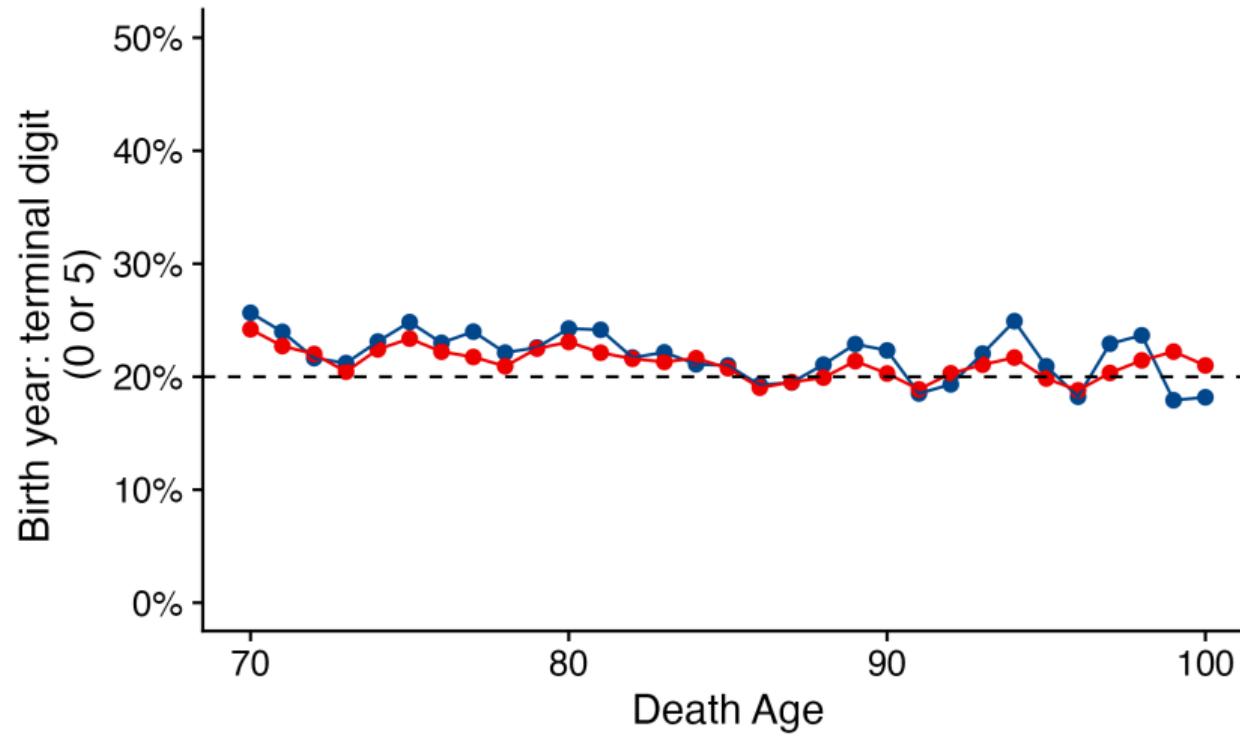
Contact: ✉ casey.breen@demography.ox.ac.uk

Mortality Crossovers (Men)
Pooled cohorts of 1890–1905

Representativeness

Age heaping

Birth cohorts of 1906-1915: Gompertz Hazard Model

$$h(x) = \mathbf{a}e^{\mathbf{b}x} \tag{2}$$

▶ $h(x) =$ hazard at age x. "Force of mortality"

Birth cohorts of 1906-1915: Gompertz Hazard Model
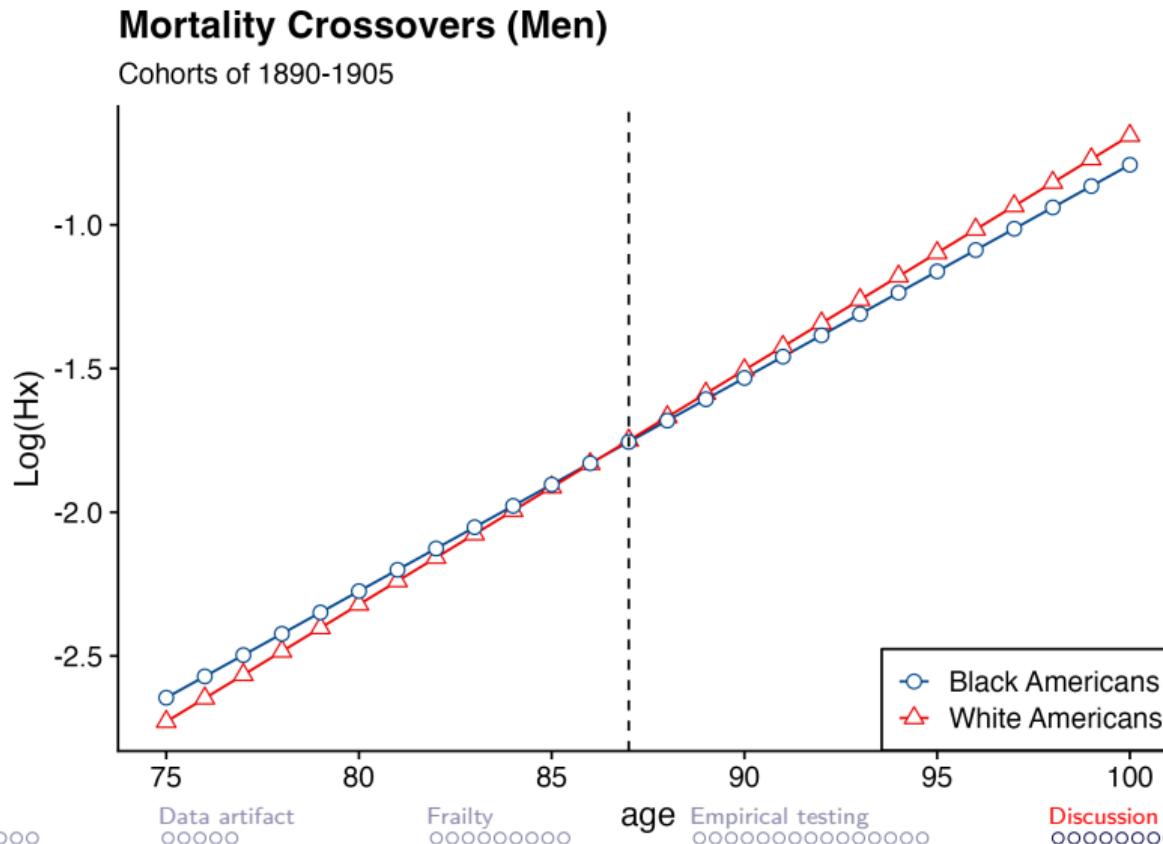
$$h(x) = \mathbf{a}e^{\mathbf{b}x} \tag{2}$$

▶ $h(x) =$ hazard at age x. "Force of mortality"

▶ **a is baseline mortality**

Birth cohorts of 1906-1915: Gompertz Hazard Model
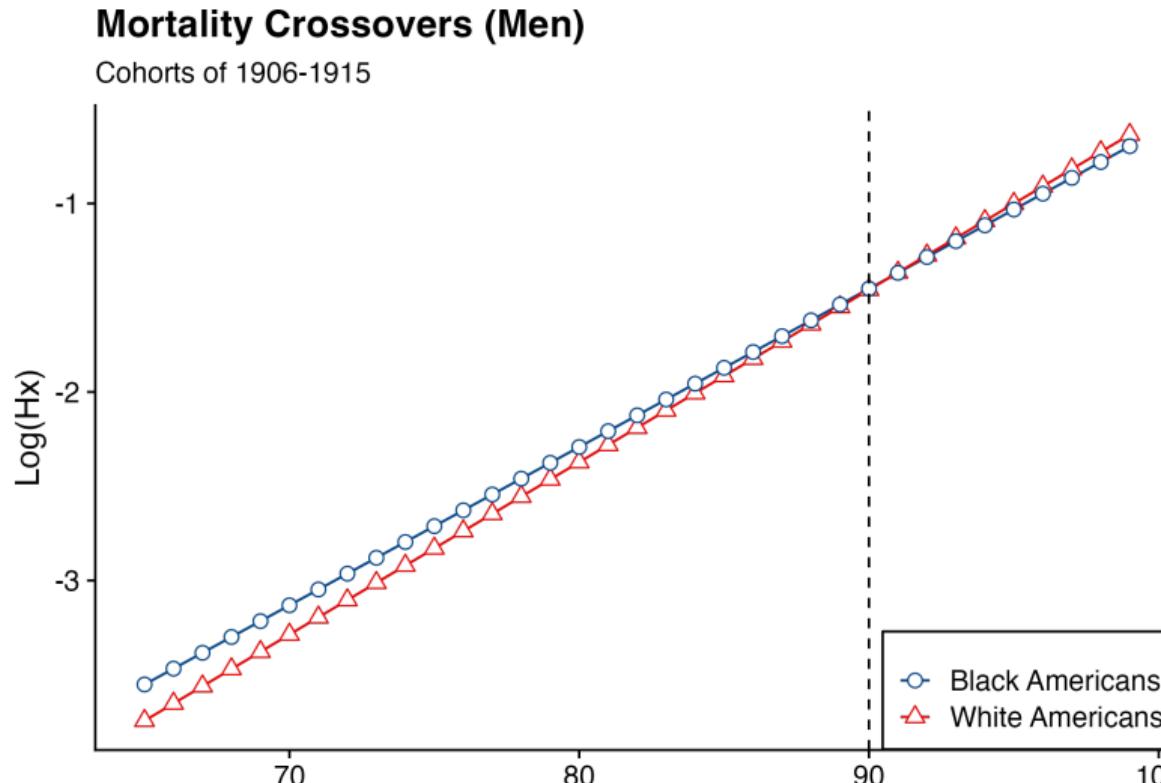
$$h(x) = \mathbf{a}e^{\mathbf{b}x} \qquad (2)$$

▶ $h(x) =$ hazard at age x. "Force of mortality"

▶ **a is baseline mortality**

▶ **b is rate of increase of mortality**
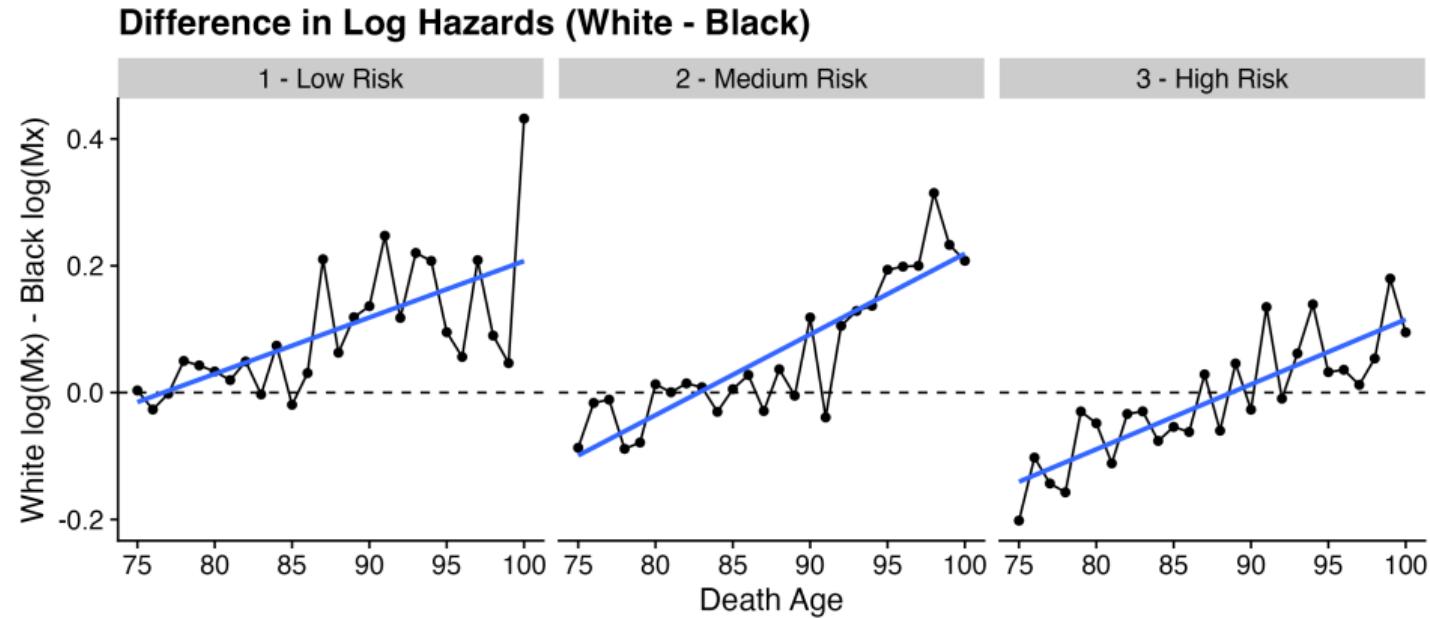
Black-white crossover (Gompertz Hazard Model)

Mortality Crossovers (Men)

Cohorts of 1890-1905

Legend:
- Black Americans
- White Americans

y-axis: Log(Hx)
x-axis: age

Black-white crossover (Gompertz Hazard Model)

Mortality Crossovers (Men)

Cohorts of 1906-1915

Stratifying on risk score

Difference in Log Hazards (White - Black)

Introduction
○○○○○○○○○○○○○○○○○○○○○○

Data artifact
○○○○○

Frailty
○○○○○○○○○

Empirical testing
○○○○○○○○○○○○○○○○○○

Discussion
○○○○○○○○○○●○○

References
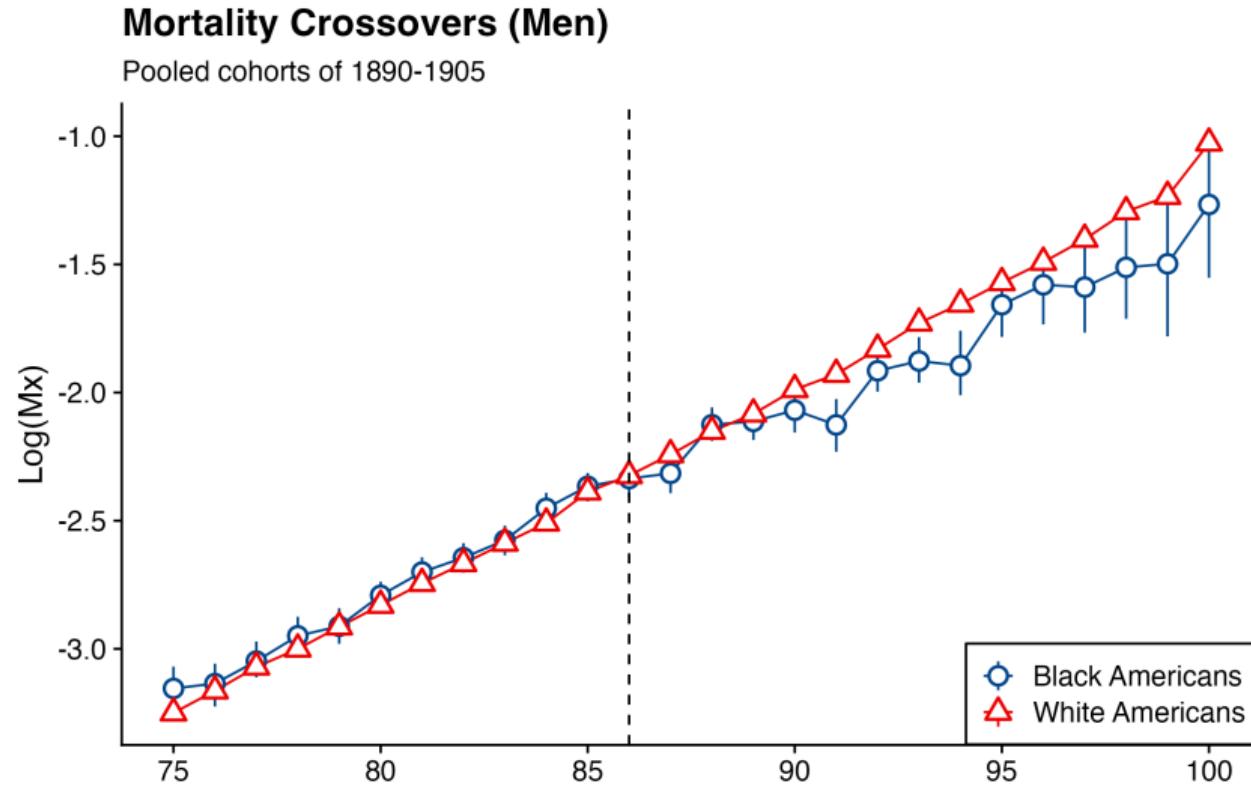
Representativeness of samples

	General Pop		CenSoc-DMF		CenSoc-DMF Siblings	
	No.	%	No.	%	No.	%
Educational Attainment						
<High School	4951782	67.3	608639	64.7	26137	66.7
High School or some college	1783203	24.3	247103	26.3	10133	25.9
Bachelors Degree	339072	4.6	48024	5.1	1664	4.2
Advanced Degree	162122	2.2	24559	2.6	820	2.1
NA	117086	1.6	12091	1.3	441	1.1
Race						
Black	656027	8.9	34159	3.6	278	0.7
Other	27778	0.4	3296	0.4	43	0.1
White	6669460	90.7	902961	96.0	38874	99.2
Marital Status						
Married	7013184	95.4	905924	96.3	38102	97.2
Not married	340081	4.6	34492	3.7	1093	2.8
Homeownership						
Homeowner	1780906	24.2	249379	26.5	11553	29.5
Not Homeowner	5572359	75.8	691037	73.5	27642	70.5
Socioeconomic Status Indicator						
Sei 1-9	1293523	17.6	138209	14.7	5513	14.1
Sei 10-14	1170543	15.9	149673	15.9	7962	20.3
Sei 15-25	1862967	25.3	246484	26.2	10028	25.6
Sei 26+	2776321	37.8	380226	40.4	14745	37.6
NA	249911	3.4	25824	2.7	947	2.4
Rural						
Rural	3183160	43.3	397739	42.3	19754	50.4
Urban	4170105	56.7	542677	57.7	19441	49.6

Crossover - Restricting to Exact Age

Mortality Crossovers (Men)

Pooled cohorts of 1890-1905

References

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

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.

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.

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.

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.

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.

Preston, Samuel H. 1996. "Population Studies of Mortality." *Population Studies* 50(3):525–536.

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.

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.

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.