The Racial Convergence of Mortality

James R. Walker

University of Wisconsin-Madison

June 10, 2010

Outline

1. Background and Summary
   - Empirical regularities
2. Frailty
   - Definition
   - Survivorship bias
   - Simple Example - Gamma Heterogeneity
3. Estimation strategy
4. Estimation Results
   - Results Period Life Tables 1900, 1910, 1920

Racial Convergence/Crossover

- Empirical regularity found in survey, census, administrative data.
- Mortality rates of blacks higher than whites at young ages, reverses at older ages.
- Age of crossover typically found to range from ages 75 to 90.

Racial Convergence/Crossover

- Empirical regularity found in survey, census, administrative data.
- Mortality rates of blacks higher than whites at young ages, reverses at older ages.
- Age of crossover typically found to range from ages 75 to 90.

Racial Convergence/Crossover

- Empirical regularity found in survey, census, administrative data.
- Mortality rates of blacks higher than whites at young ages, reverses at older ages.
- Age of crossover typically found to range from ages 75 to 90.

Why the interest?

- Not discussed explicitly in papers.
- Presumption - is convergence an artifact of the data or it is real?
- If real, how to interpret?

Explanations

- Artifact (Preston et al.) due to measurement error in age of death of blacks. Blacks systematically over report age at death (2 years).
- Frailty (Manton et al.) **Unobserved** (time invariant) differences leads to selected sample. The more frail blacks die at younger ages, so on average blacks at later ages are more robust.
- Compositional (Elwert et al.) **Observed** (and perhaps time varying) covariates of blacks and whites have different distributions (e.g., education, region of residence).
- True. Race-specific mortality schedules differ and cross.

Explanations

- Artifact (Preston et al.) due to measurement error in age of death of blacks. Blacks systematically over report age at death (2 years).
- Frailty (Manton et al.) **Unobserved** (time invariant) differences leads to selected sample. The more frail blacks die at younger ages, so on average blacks at later ages are more robust.
- Compositional (Elwert et al.) **Observed** (and perhaps time varying) covariates of blacks and whites have different distributions (e.g., education, region of residence).
- True. Race-specific mortality schedules differ and cross.

Explanations

- Artifact (Preston et al.) due to measurement error in age of death of blacks. Blacks systematically over report age at death (2 years).
- Frailty (Manton et al.) **Unobserved** (time invariant) differences leads to selected sample. The more frail blacks die at younger ages, so on average blacks at later ages are more robust.
- Compositional (Elwert et al.) **Observed** (and perhaps time varying) covariates of blacks and whites have different distributions (e.g., education, region of residence).
- True. Race-specific mortality schedules differ and cross.

Explanations

- Artifact (Preston et al.) due to measurement error in age of death of blacks. Blacks systematically over report age at death (2 years).
- Frailty (Manton et al.) **Unobserved** (time invariant) differences leads to selected sample. The more frail blacks die at younger ages, so on average blacks at later ages are more robust.
- Compositional (Elwert et al.) **Observed** (and perhaps time varying) covariates of blacks and whites have different distributions (e.g., education, region of residence).
- True. Race-specific mortality schedules differ and cross.

Key Conclusions

- The literature too quick to dismiss frailty.
- Common to neglect left-truncation in demography literature. Can't fathom why.
- Seem to forgotten Sheps and Menken (1973) and work of Manton, Vaupel and coauthors mid- to late 1980s.
- I have drunk the Kool-Aid. Results (if believable) imply frailty operates very early in life.

Key Conclusions

- The literature too quick to dismiss frailty.
- Common to neglect left-truncation in demography literature. Can't fathom why.
- Seem to forgotten Sheps and Menken (1973) and work of Manton, Vaupel and coauthors mid- to late 1980s.
- I have drunk the Kool-Aid. Results (if believable) imply frailty operates very early in life.

Key Conclusions

- The literature too quick to dismiss frailty.
- Common to neglect left-truncation in demography literature. Can't fathom why.
- Seem to forgotten Sheps and Menken (1973) and work of Manton, Vaupel and coauthors mid- to late 1980s.
- I have drunk the Kool-Aid. Results (if believable) imply frailty operates very early in life.

Key Conclusions

- The literature too quick to dismiss frailty.
- Common to neglect left-truncation in demography literature. Can't fathom why.
- Seem to forgotten Sheps and Menken (1973) and work of Manton, Vaupel and coauthors mid- to late 1980s.
- I have drunk the Kool-Aid. Results (if believable) imply frailty operates very early in life.

Frailty

- Name implies "biological" process.
- Anything not measured. Can be biological, social and or behavioral.
- Unobserved, simplicity assume time invariant.
- If frailty health environment or disease risk, expect that to vary over time.
- Time-varying frailty research frontier, but requires some data to be estimable.

Survivorship bias

- To be observed, an individual must survive to the beginning of the observational (sampling) period.
- Failue to recognize left-truncation yields valid statistical inference only if mortality follows an exponential distribution.
- Survivorship bias increases with the length of truncation.
- Not recognizing left-truncation may be one reason for the observed variability of the cross over ages.

Studies Neglecting left truncation

- Dupre, M. et al. (2006) Demography
- Hayward and Gorman (2004) Demography
- Hayward et al. (1997) Journal of Health Behavior
- Johnson (2000) Demography

Inferential Problems

- Inference: Absent controls, ambiguous what is population of interest. Cohort or those who survive to observation age age?
- Concern: authors believe making inference on cohort, but inference applies to survivor population.
- If cohort survivor density is $f(t)$ and minimum observed age is τ, sample density is $g(t) = f(t|t \geq \tau)$.
- Can show that a variable, X, may have a large effect on cohort survivorship, but estimated effect in sample may be attenuated.
- Because X may also effect who survives to τ.

Truncated Sample has less information

- Truncated sample less information than (common) censored sample.
- Censored sample provides information on the fraction of censored observations.
- Example: Top coded income. The random variable income is truncation, but sample is censored; sample provides information on proportion topcoded.
- Truncated sample: all respondents satisfy selection criterion, need external information to determine extent of trucation.

Truncated Sample has less information

- Truncated sample less information than (common) censored sample.
- Censored sample provides information on the fraction of censored observations.
- Example: Top coded income. The random variable income is truncation, but sample is censored; sample provides information on proportion topcoded.
- Truncated sample: all respondents satisfy selection criterion, need external information to determine extent of trucation.

Truncated Sample has less information

- Truncated sample less information than (common) censored sample.
- Censored sample provides information on the fraction of censored observations.
- Example: Top coded income. The random variable income is truncation, but sample is censored; sample provides information on proportion topcoded.
- Truncated sample: all respondents satisfy selection criterion, need external information to determine extent of trucation.

Truncated Sample has less information

- Truncated sample less information than (common) censored sample.
- Censored sample provides information on the fraction of censored observations.
- Example: Top coded income. The random variable income is truncation, but sample is censored; sample provides information on proportion topcoded.
- Truncated sample: all respondents satisfy selection criterion, need external information to determine extent of trucation.

Minimum recognition of Left Truncation

- Population density of survivorship: $f(t) = h(t)S(t)$.
- Let $m(t)$ denote sample density of survivorship.
- $m(t) = \frac{f(t)}{\Pr(T \geq \tau)} = \frac{h(t)S(t)}{S(\tau)} = h(t)S(t-\tau)$
- At a minimum, authors should include age t and duration since start of sample $t - \tau$ in their analyses.
- Johson includes only age. Same for Hayward et al. (1997). Hayward Gorman (2004) includes age and two cohort dummies (not quite enough).

Simple Example

Gamma Heterogeneity within proportional hazard

- Hazard: $h(t|\theta) = h_0(t)\theta, \qquad \theta \sim \Gamma(a,b) \quad f(\theta) = \frac{\theta^{a-1}e^{-\theta/b}}{\Gamma(a)b^a}.$
- Integrated hazard: $H_0(t) = \int_0^t h_0(u)\,du.$
- Parameterize $(a = b^{-1})$, $E[\theta] = 1$ and Var(θ)=b.
- Population hazard

$$
\begin{aligned}
S_p(t) &= \int_0^\infty S(t|\theta)\,d\Gamma(\theta) \\
&= (1 - b\ln S_0(t)^{-1/b} \\
&= (1 + bH_o(t))^{-1/b}
\end{aligned}
$$

Updating

Example continued

Apply Bayes' theorem and can show that:

$$f(\theta|t) \sim \Gamma(a = 1, \kappa), \quad \kappa = b/(1 + H_0(t))$$

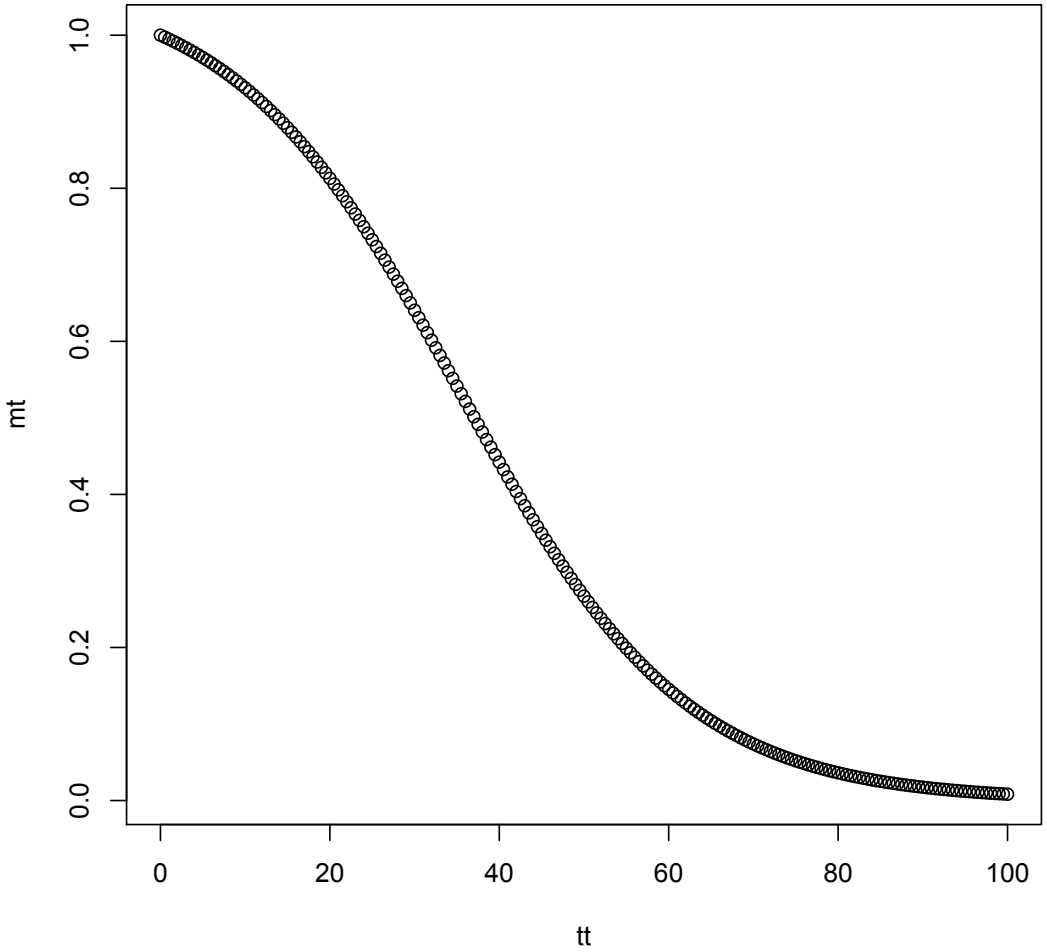
Hence,

$$\mu(t) = E[\theta|T \geq t] = \frac{b}{1 + bH_o(t)}$$

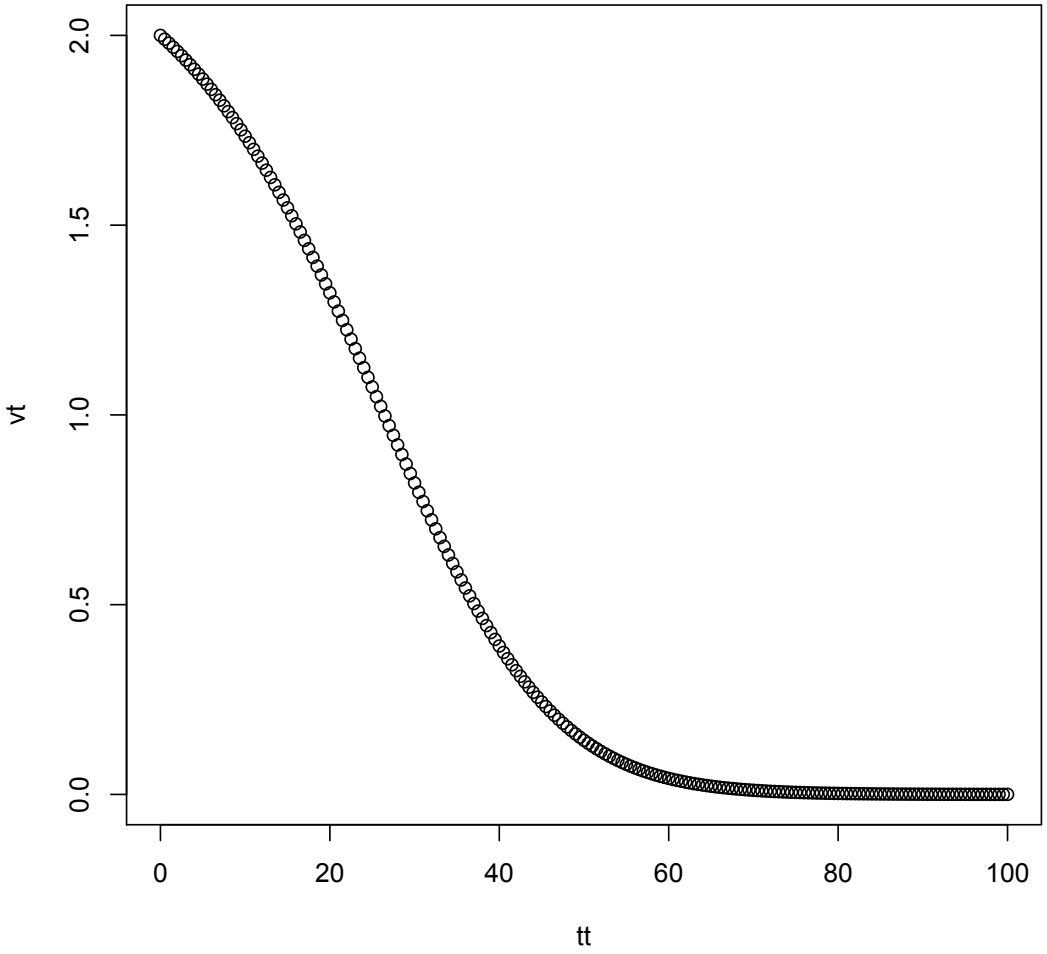
$$\sigma^2(t) = Var(\theta|T \geq t) = \frac{b^2}{1 + H_0(t)}$$

See that as $t \to \infty \quad \mu(t), \sigma^2(t) \to 0$.

$\mu(t)$



$\sigma^2(t)$



Empirical Evidence on Frailty

- Elo and Preston (1996) and Johnson (2000) comment that their samples provide no evidence in support of frailty.
- Johnson (2000) analyzes AHEAD data, $\tau \geq 70$.
- Elo and Preston (1996) analyze National Longitudinal Mortality Study. Follow respondents of CPS for ten years. Yet, likely those who die are concentrated in older ages, and survived to reasonably elderly age.
- Have the NLMS and will investigate.

Estimating magnitude of survivor bias

- Hayward et al. (1997), Mare and Palloni (1986), Mare (1991) have used the NLS-Older Men Cohort.
- NLS-OM born 1906-1921.
- By nature of truncated sample, must look beyond sample information.
- Follow cohort approach, augment sample information with external cohort information.
- US Commerce Department, NCHS, Mortality Project (UC Berkeley) sources of information.

Period to Cohort Lifetables

- Use set of period lifetables to build (synthetic) cohort lifetable.
- Mortality Project - cohort life table for US males 1933 birth cohort (not by race).
- Use Commerce Department period life tables by race.
- Ideal is single year period table for each year.
- Available for later years, estimate synthetic cohort and compare with national samples (HRS, Medicare, etc.).

NLS-Old Men

- NLS-OM fielded in 1966 through 1981 and in 1990 (respondent or partner, "exit interview").
- Matched with National Death Index.
- Covers mortality process for 24 years (NDI match eliminates attrition problems).
- Has some information on parents of respondents. Compare characteristics to 1920 Census.

Relax Gamma distribution

- Relax the assumption $\theta \sim \Gamma(a, b)$
- Assume $\theta \sim G(\theta) \quad E[\theta] = 1, \quad Var(\theta) = \sigma^2 < \infty.$
- Population survivor becomes:

Population Survivor

$$
\begin{aligned}
S_p(t) &= \int_0^\infty S(t|\theta)\, dG(\theta) \\
&= \int_0^\theta \exp\left[-\theta H_0(t)\right] dG(\theta) \\
&= E_\theta \left[e^{-\theta H_0(t)} \right] \\
&= \mathscr{L}(H_0(t)).
\end{aligned}
$$

- $\mathscr{L}(\cdot)$ is the Laplace transform.

Laplace Transform

- Let $z = H_0(t)$.
- $\mathscr{L}(z) = E\left[e^{-\theta z}\right]$.
- $E\left[\theta^k\right] = (-1)^k \frac{d^k}{dz^k}\mathscr{L}(z)|_{z=0}$
- Thus, for time rescaled by $H_0(t)$, $S_p(t)$ maps into moments of $G(\theta)$.

Parametric Estimation Problem

- Need to parameterize $H_o(t)$;
- May specify $G(\theta)$ within a class of distributions (e.g., Power Variance Transform which includes gamma, inverse Gaussian and compound Poisson);
- Or follow Heckman and Singer (1984) and specify $G(\theta)$ nonparametrically as discrete distribution: $(\theta_1, \ldots, \theta_K, \pi_1, \ldots, \pi_K; \sum_{j=1}^{K} \pi_j = 1)$. Where K is a parameter to be estimated as well.
- Data: population survivor function and then by analogy principle can estimator parameters of $H_o(t)$ and $G(\theta)$ that minimize the distance between the observed (population survivor function) and population survivor function predicted by the model.

Objective Functions

- Two estimation procedures (based on available software)
  - Simulate histories from observed $S_p(t)$ estimate by MLE.
  - Minimum χ^2 estimation: $min_\varphi \sum_{t=1}^{B} \frac{\left[S_p(t)-S(t|\varphi)\right]^2}{S(t|\varphi)}$. (Under development).

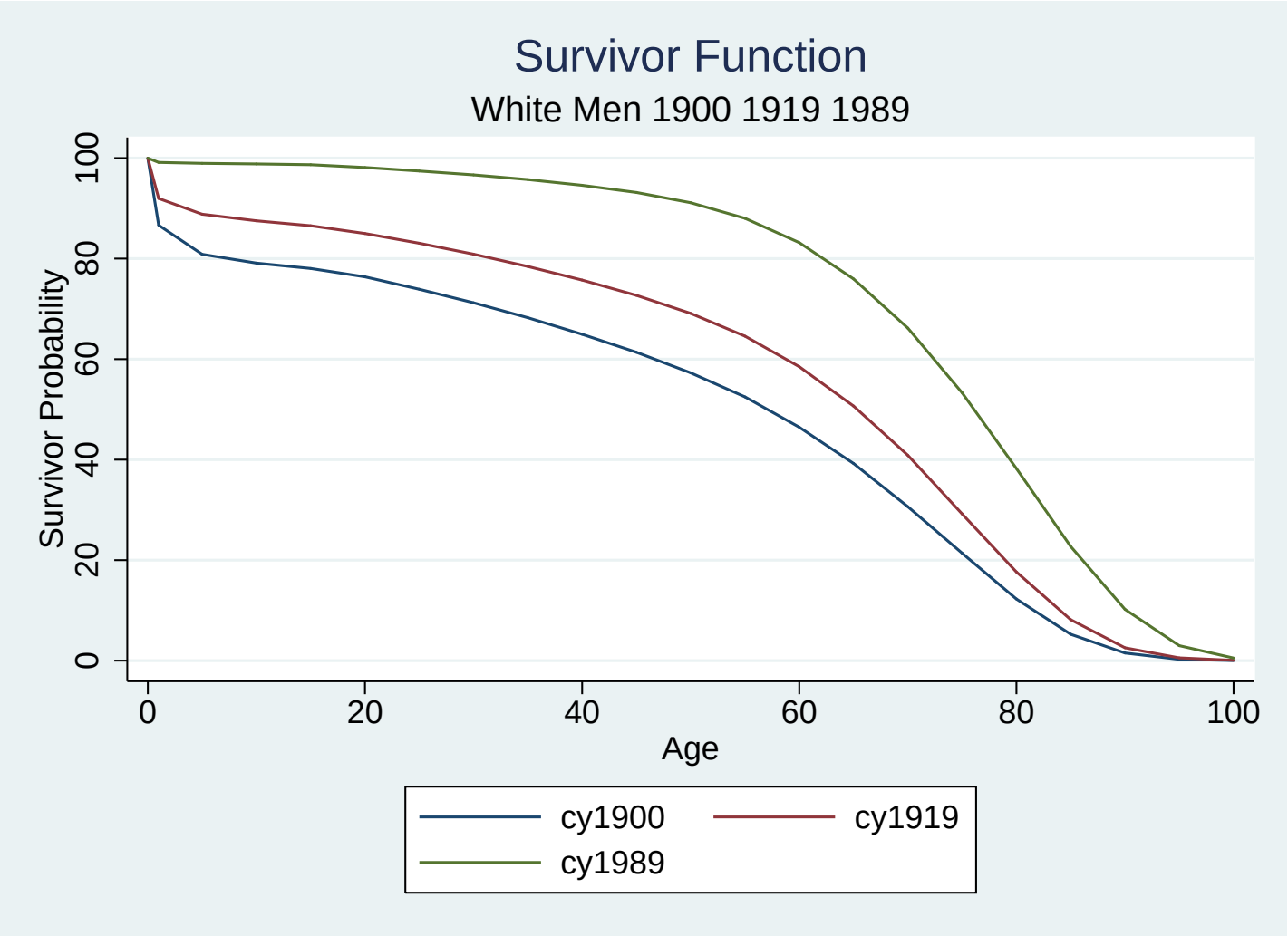
Period Life Table 1900 by Race

Table 1: Period Life Table 1901-1910

Statistic	Blacks	Whites
IMR (per 1,000)	241.4	127.4
$\Pr(T > 45)$	.39	.63
$\Pr(T > 65)$	.18	.40
$\Pr(T > 70)$	.12	.31
e_0	32.6	49.3

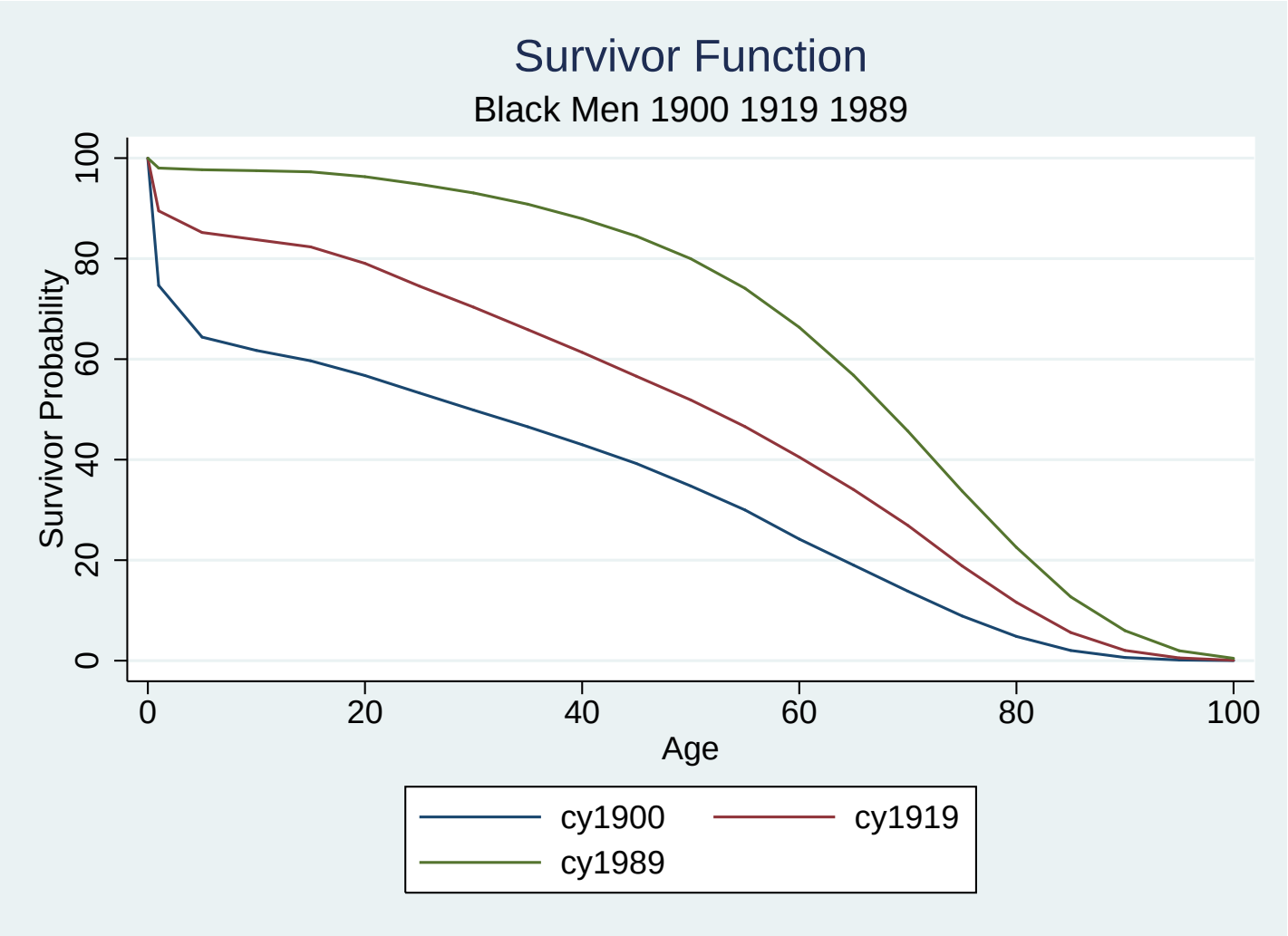
Survivor Function Cohorts 1900, 1910, 1920

White Men

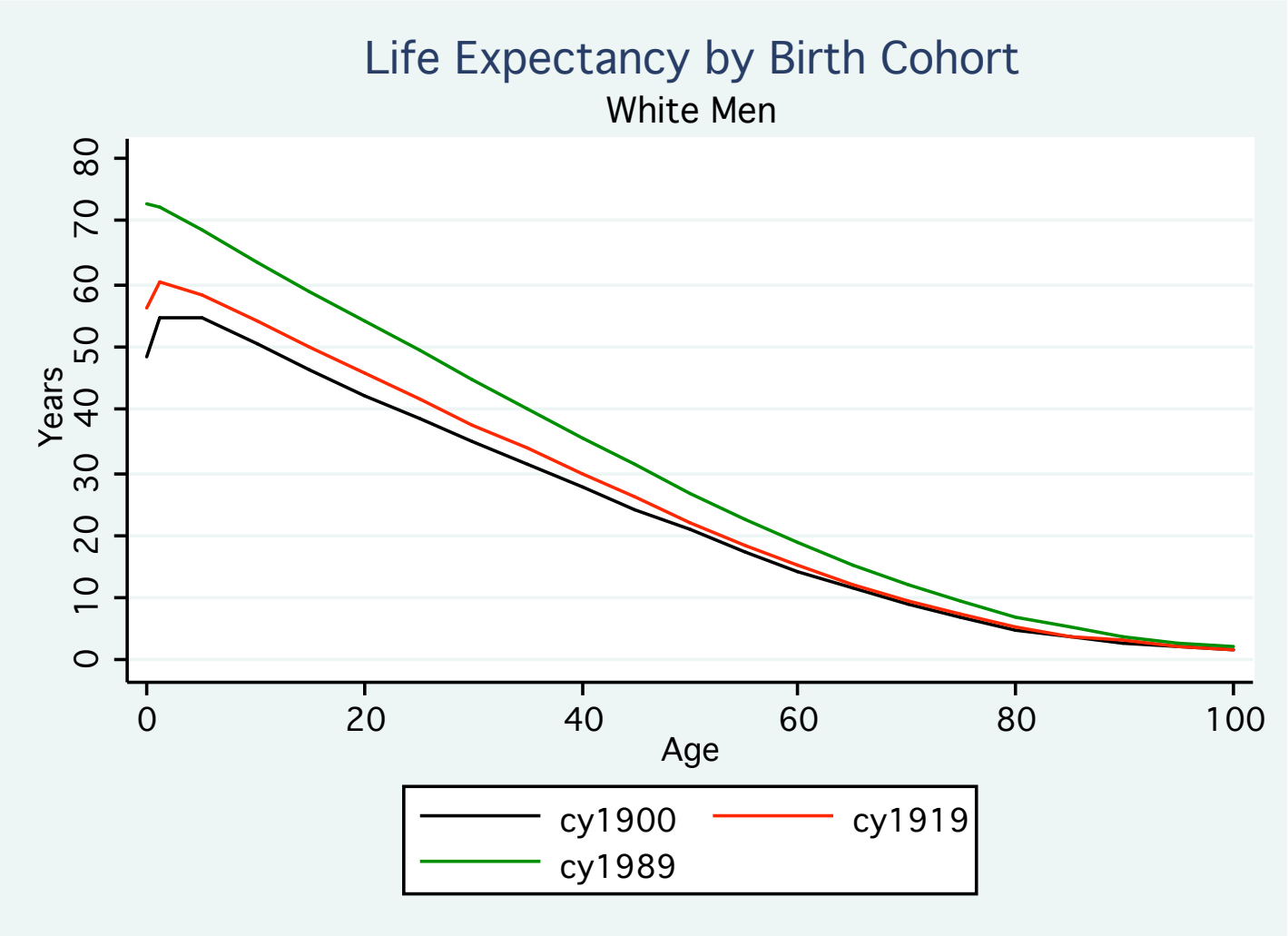


Survivor Function Cohorts 1900, 1910, 1920

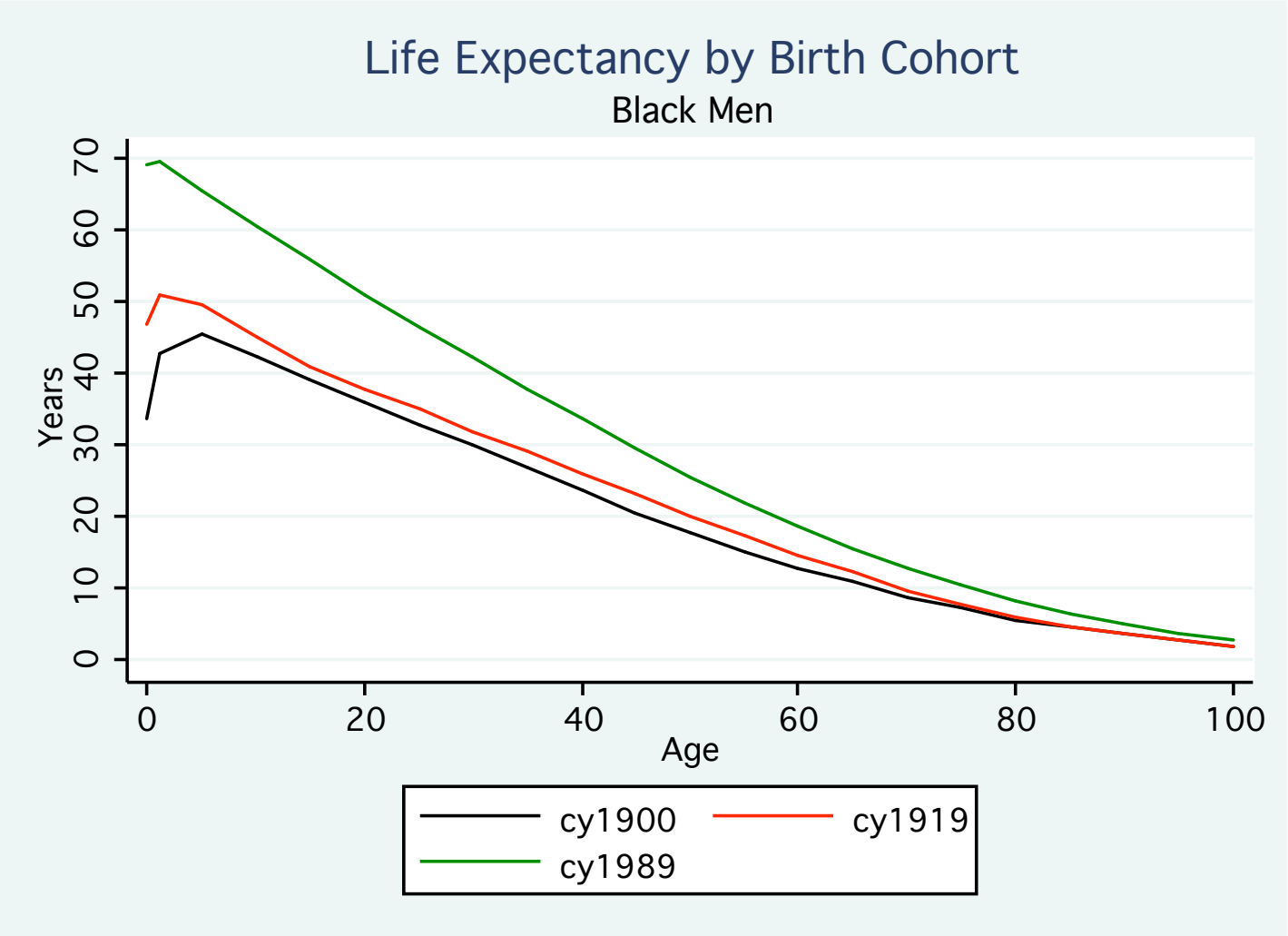
Black Men



Life Expectancy: Whites 1900-1990



Life Expectancy: Blacks 1900-1990



Gompertz Model Estimates

Full and Truncated Samples

	Full		Trunc	
Parm	Est	SE	Est	SE
γ_0	.1505	.0316	-.5984	.0699
γ_1	2.044	.0182	3.748	.1011
β	-.1403	.0235	-.0869	.0358
loglike	-10795.6		-2322.9	
N	15,000		10,398	
Mean completed spell	115.28		27.57	

Note: $h(t) = exp(\gamma_0 + \gamma_1(t-1) + \beta(cohort/10))$
Time measured in units of half-years, for estimated parameters reflect time scaled by 0.01.

Gompertz with finite mixing distribution

Full and Truncated Samples

	Full		Trunc	
Parm	Est	SE	Est	SE
γ_0	7.4712	.0417	-.5980	NaN
γ_1	3.326	.0253	3.747	.3090
β	-.1166	.0208	-.0870	.0382
c	-7.836	.0381	-.000029	NaN
π_1	.1455	.0029	.0919	NaN
loglike	-4076.5		-2322.9	
Hess #	1		Singular	

Mixing distribution: $\theta_1 = 0, \quad \theta_2 = 1$. c is factor loading and $|c| = \sigma(\theta)$.
Models with three points of support is singular in full sample.

Quadratic Hazard Models

Full Sample

	1 point		2 points	
Parm	Est	SE	Est	SE
γ_0	-.5991	.0253	-.3091	.0354
γ_1	-3.438	.0461	-5.082	.0603
γ_2	5.470	.0468	8.090	.0751
β	-.1591	.0164	-.2340	.0212
c	—		-2.1141	.0407
π_1	—		.7553	.0064
Loglike	-6882.9		-6337.6	
Hess #	1		1	

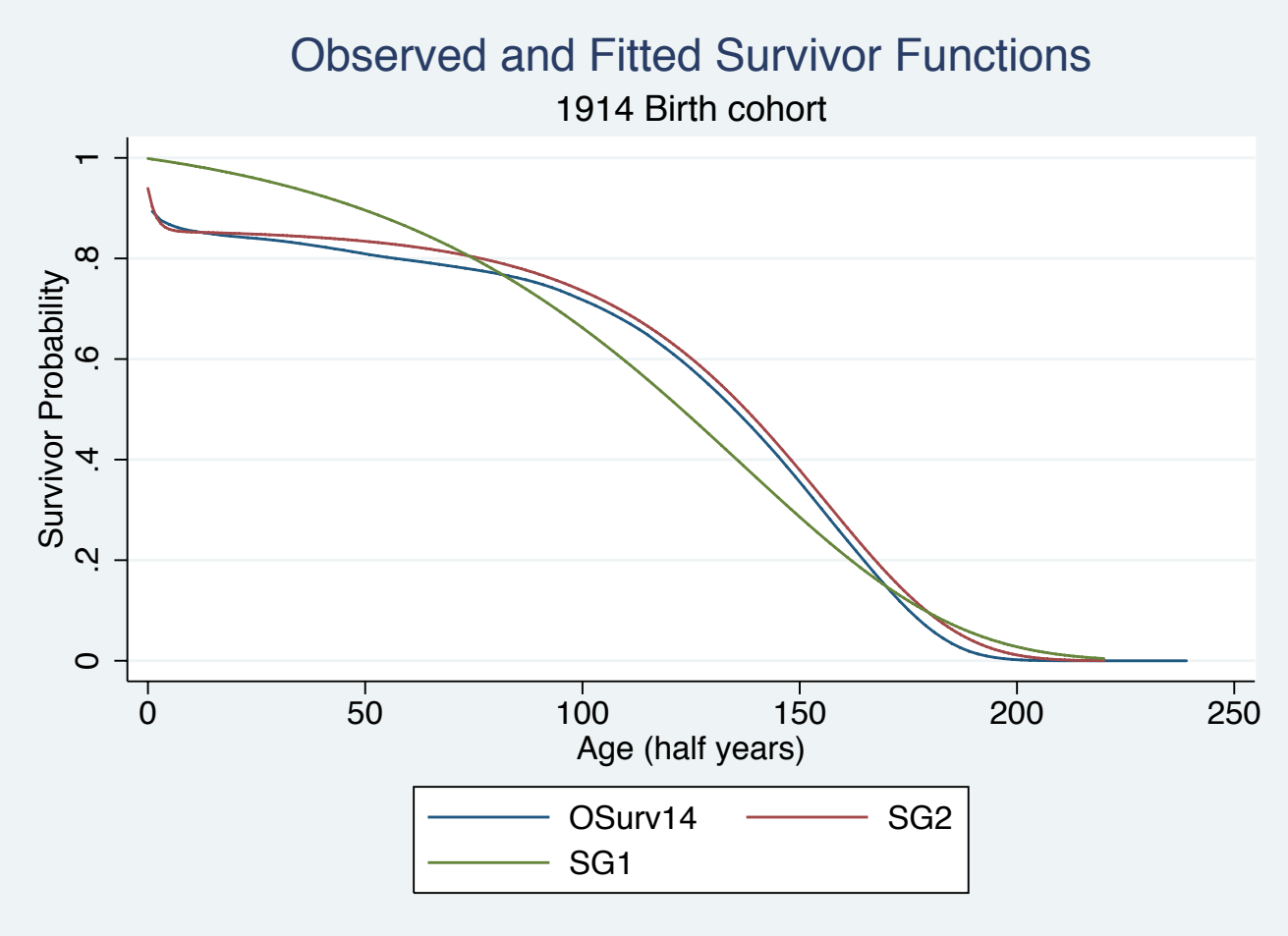
$$h(t|\theta) = exp(\gamma_0 + \gamma_1(t-1) + \gamma_2(t^2-1)/2 + \beta(cohort/10) + c\theta)$$

Time measured in half-years; time scaled by 0.01.

Model with three support points is singular.

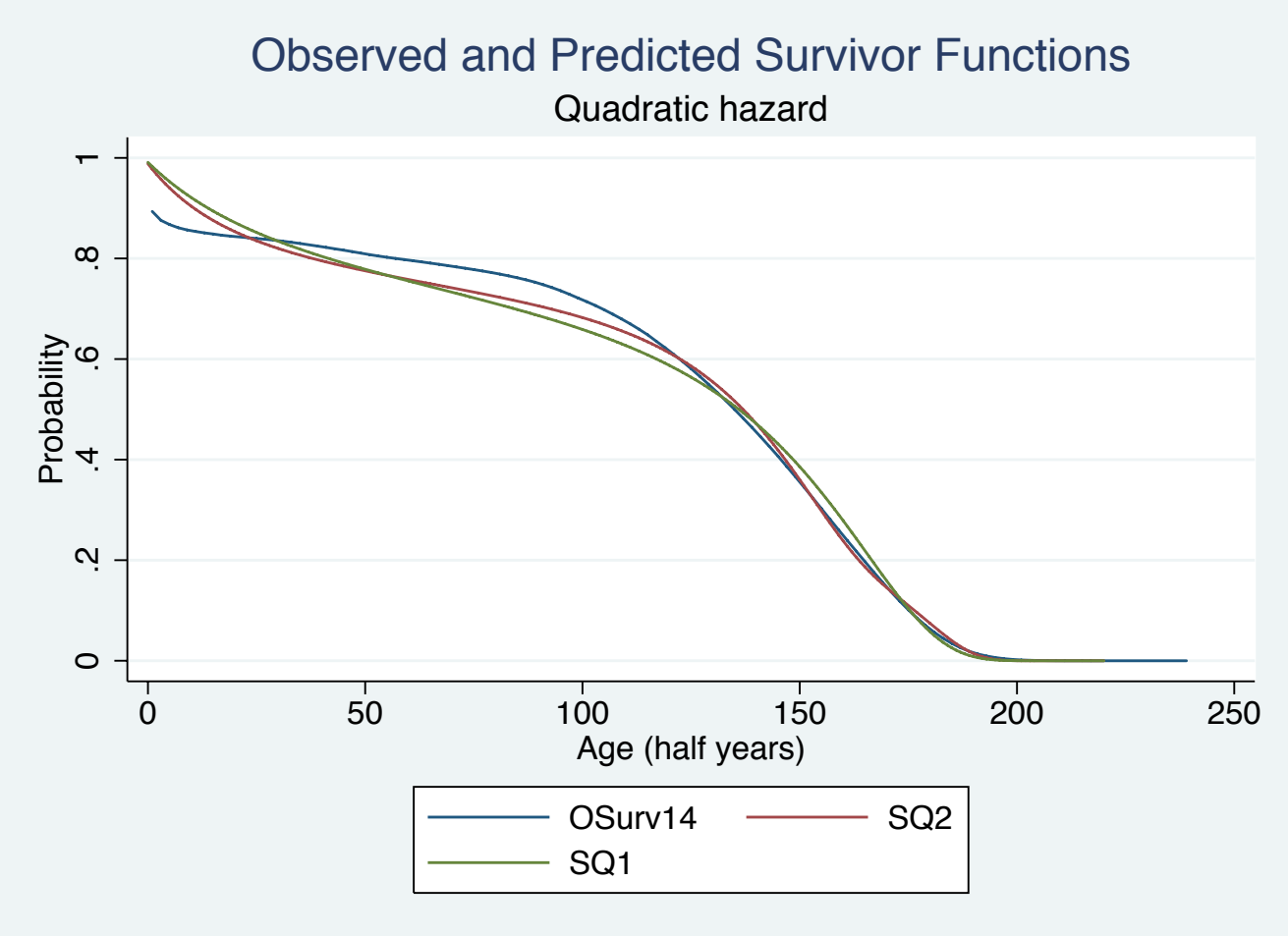
Gompertz Models - Estimated Versus Observed

Survivor Functions: 1914 Cohort

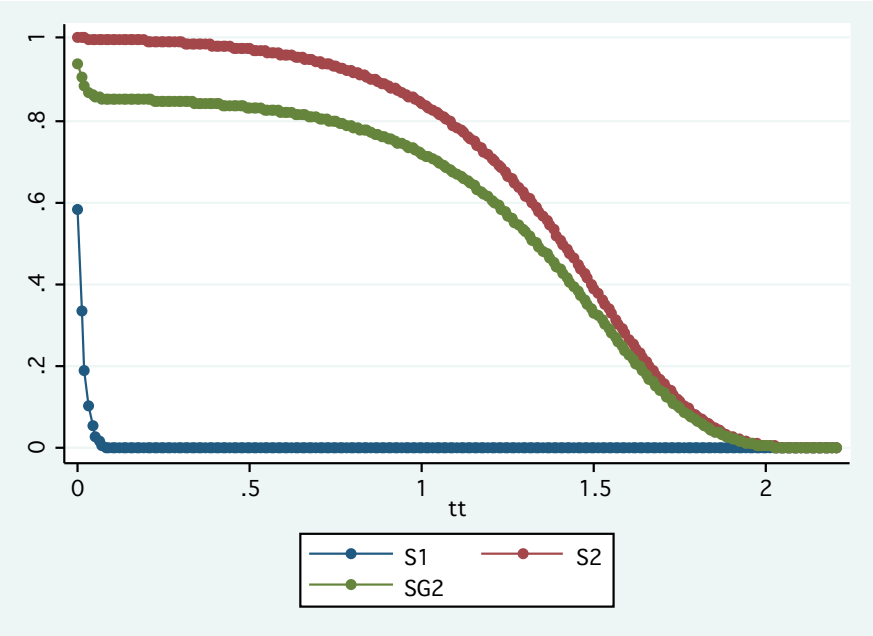


Quadratic Models - Estimated v Observed

Survivor Functions: 1914 Cohort



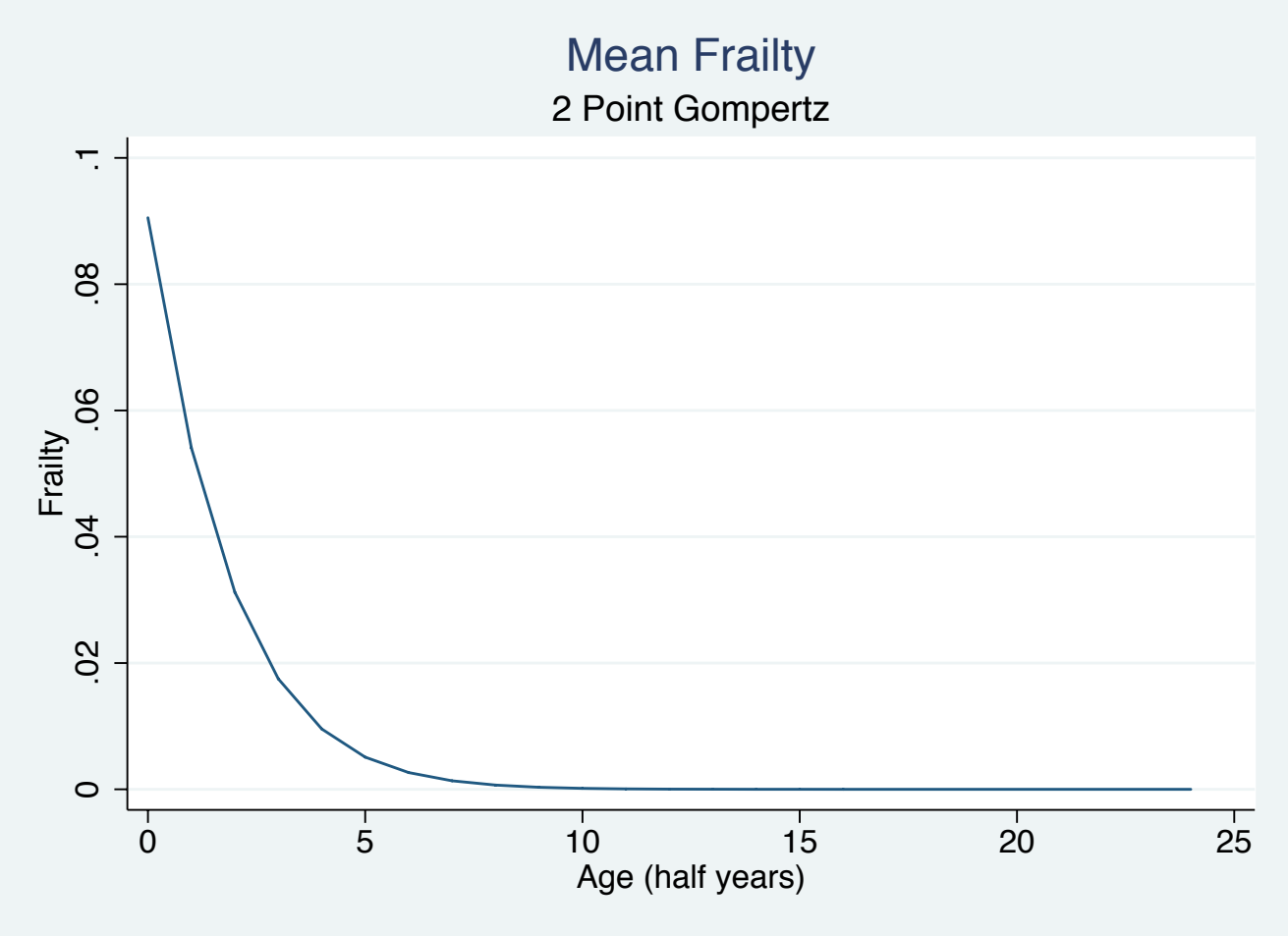
Gompertz conditional survivor function



$E[T|\theta = 0] = .67$ years
$E[T|\theta = 1] = 70.0$ years,
$E[T] = 59$ (model), sample $\bar{T} = 57.8$ years.

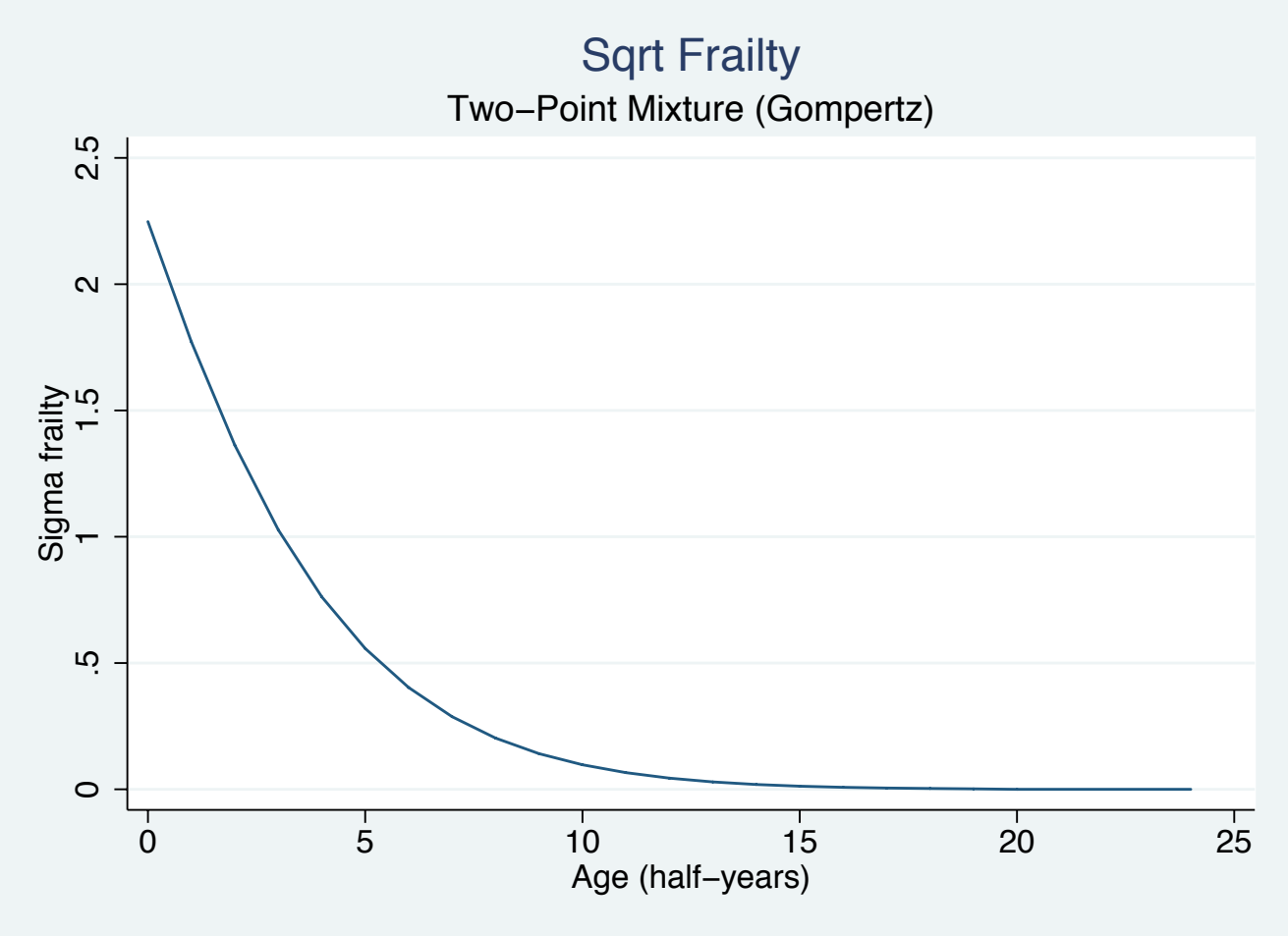
Mean Frailty by Age

Two-point Gompertz



Sigma Frailty by Age

Two-point Gompertz



Future Work

- Pilot analysis suggests full analysis will be fruitful.
- Need to develop life table by race for NLS-OM cohorts.
- Develop min χ^2 software.
- Data analysis compare characteristics of parents of NLS-OM respondents with information recorded in 1920 census.