Journal of Econometrics 50 (1991) 183–203. North-Holland

Black–white mortality inequalities*

Jere R. Behrman
University of Pennsylvania, Philadelphia, PA 19104, USA

Robin Sickles
Rice University, Houston, TX 77251, USA

Paul Taubman
University of Pennsylvania, Philadelphia, PA 19104, USA

Abdo Yazbeck
Texas A&M University, College Station, TX 77843, USA

Racial inequality, particularly between blacks and whites, long has been of major concern in the United States. This inequality may take a number of forms, for example, with regard to schooling, housing, health, employment options, and income. In this paper we estimate how much of the 'observed' racial inequality in age-specific death rates in the Retirement History Survey can be attributed to observed differences in variables in hazard functions. The exact answer depends on what independent variables are excluded, but fairly standard explanatory variables explain 50% to much more of the observed greater death rates of black men.

1. Introduction

Racial inequality, particularly between blacks and whites, long has been of major concern in the United States. This inequality may take a number of forms, for example, with regard to schooling, housing, health, employment options, and income.[1] The National Academy of Science report on *Blacks and American Society* edited by Jaynes and Williams (1989) reviews the

*This research was supported by a National Institute on Aging Grant 1-R01-AG-05384-01. Computer services were made available to us through the Houston Area Research Center's NEC/SX-2 supercomputer, one of the world's fastest uniprocessors. The authors thank James Heckman and James Walker for making available to us an early version of their CTM software [Yi et al. (1986)]. The authors also thank Esfandiar Maasoumi and an anonymous referee for comments that strengthened our paper. The usual caveat applies.

0304-4076/91/$03.50

recent status of black Americans. Jaynes and Williams (p. 6) summarize their main findings on the status of blacks in America in the late 1980's succinctly:

- 'By almost all aggregate statistical measures – incomes and living standards; health and life expectancy; educational, occupational, and residential opportunities; political and social participation – the well-being of both blacks and whites has advanced greatly over the past five decades.
- By almost all the same indicators, blacks remain substantially behind whites.'

Economists concerned with racial inequalities in the United States have concentrated on the nature of such inequalities in income, though other differences also have been examined.[1] A major focus has been to try to understand to what extent such inequalities have been due to average differences in observed characteristics thought to underlie an outcome (e.g., in the schooling-underlying wage rates) and to what extent they are due to differences in the effects of those characteristics. Or, to put the question slightly differently, how much of the existing black–white differences in an outcome of interest would disappear if blacks had the same observed characteristics as do whites?

Our concern in this paper is to explore black–white inequalities in the United States with regard to mortality for older men and to what extent such differences are associated with observed differences in characteristics such as marital status, education, and occupation.

Throughout much of their lifespan, blacks have a higher age-specific death rate than whites in the United States (though there may be a cross-over at later ages). For example, the annual death rate in 1960 for males aged 50 was about 9.5 and 15.6 per 1000 for whites and blacks, respectively. Kitigawa and Hauser (1973, p. 103) indicate that at about the same time, the remaining life expectancy at age 55 for white and black males was 19.5 and 18.4 years, respectively. Jaynes and Williams (1989, p. 427) report that between 1900 and 1984 the expected remaining years of life at age 65 increased from 11.5 to 14.8 for white men and from 10.4 to 13.4 for black men. Also, as shown in fig.

[1]For example, see Shulman (1987), Kahn and Sherer (1988), Andrisani (1977), Welch (1973), Smith (1984), Orazem (1987), Darity (1982), Ashenfelter (1977), Freeman (1973), Smith and Welch (1977), and Welch (1974). For a theoretical treatment of inequality measurement see Maasoumi (1986). The approach that we adopt in this paper is less formal and more descriptive than the generalized entropy inequality measures exposited by Maasoumi and applied elsewhere [Maasoumi and Nickelsburg (1988)]. However, as one important aspect of inequality, our estimates of mortality differences by race could be used to modify and, we think improve, existing generalized entropy inequality estimates.

1 and table 4, the death hazard rate is much higher for blacks than whites in the years covered in the Retirement History Survey (RHS), which is used in this paper.

Over the life cycle, whites and blacks face substantially different environments and have major differences in education, earnings, occupation, and marital status, all of which variables have been found to be related to morbidity and mortality in studies by Behrman, Sickles, and Taubman (1988), Sickles and Taubman (1986), Rosen and Taubman (1982), Kitigawa and Hauser (1973), and others. While such observed characteristics may account substantially for black–white mortality differentials, there also may be major causes that are not observed in most socioeconomic data sets. Jaynes and Williams (1989, p. 425), for example, suggest that such factors may be quite important: 'Black adults reach age 65 with life histories of disproportionate prevalence of acute and chronic disease, illness, and disability. They have had poorer quality of health care from conception and birth, continuing exposure to greater and more severe environmental risk factors, and the stress of prejudice and discrimination [Cooper et al. (1981)]. Cohort data for cause-specific mortality and morbidity over the past four decades suggest the presence of accumulated deficits across the early years of the life course. These deficits place black older people at greater risk for morbidity and mortality than whites of comparable ages.'

In this paper, we use the males in the RHS to see how much of the observed inequality in mortality hazard rates (the age-specific death rate in a year t divided by the survivors in that age cohort up to time t) is eliminated once we control for certain observed variables. We do the analysis both with several accelerated time-to-failure models and proportional hazard models with frailty differences among individuals modeled parametrically and non-parametrically. We estimate separate equations for blacks and whites and test whether the data should be pooled. We also evaluate the effect of differences in mean values for the right-side variables across these different specifications.

We find that a number of socioeconomic variables, including marital status, income, and education, are significantly associated with mortality for one or both races. We reject the null hypothesis that the coefficients of an estimated equation are the same for the two races. But we find that most of the observed inequalities in death hazards is consistent with the difference in the mean values of the observed socioeconomic variables rather than reflecting different coefficients due to factors such as those suggested by Jaynes and Williams.

In the following sections of the paper, we discuss the sample used in our analyses (section 2), present an overview of the framework on which our empirical analyses is based (section 3), introduce the exhaustive set of statistical specifications considered in our empirical work (section 4), review our estimation results (section 5), and offer concluding remarks (section 6).

2. The data

The Retirement History Survey (RHS) was started in 1969 with about 11,000 men and women, though we restrict our analysis to men to reduce heterogeneity. At that time it was a nationwide random sample of heads of households aged 58–63. The sample members were reinterviewed every two years through 1979. We have constructed a longitudinal file from the interviews through 1977, so we include ages from 58 through 73 in our analysis. Death information has been collected from two sources. The RHS records death as a reason for non-reinterview if this is known to be the case, generally through interviewing the surviving spouse. This source is incomplete. The other source is the Social Security files, which record deaths reported to the Social Security Administration by month and year as part of the process of issuing benefits as a result of death (such as burial grants and survivor benefits for dependent children) and making necessary adjustments in old age benefits. We currently have this death information through 1977 (with incomplete data into 1979) by which time about 22 percent of the white and 27 percent of the black men had died. We have compared these two sources and there are only two cases of death recorded in the RHS but not in the Social Security files. Moreover, the Social Security files' dates of death are in accord with RHS in that the individual does not give interviews after Social Security records his or death. Duleep (1986), following up on Rosen and Taubman (1982), has indicated that the Social Security files now record nearly all deaths.

The RHS contains information on the respondents and spouses including age, education, wealth, earnings, pensions,[2] Social Security benefits, earnings covered by Social Security annually for the period 1951–1976, number of children, current and previous occupation, marital history, spouse's earnings, health status, retirement status and plans, and some aspects of life style including contact with children.

Right censoring occurs in our samples. We assume that this censoring is random in both samples. However, the samples are not random draws from the population of failure times of the cohorts born 1911 to 1916 since they

[2]Although pension income is time-varying in the RHS we treat it in our analyses as fixed. The reason is that different years' pension income figures have wide fluctuations in missing observations and deleting those missing observations and hence those individuals from the sample would have substantially reduced its size as well as increased the percentage of censored respondents. We use the level of pension income (1000$) reported by the head of household or surviving spouse in 1977 and note that we may be biasing upward the effect that pension income has on reducing the death hazard. Results without pension income are otherwise quite comparable to those presented in tables 1–3.

only contain those still alive in 1969. Inferences to a broader population of individuals therefore may not be accurate.

3. Framework

We are interested in estimating the determinants of mortality, and how those determinants differ between black and white males, conditional on a set of observed predetermined variables. The probability that a person dies within a period can be approximated by the hazard of dying, $\lambda(t)$. In a standard utility-maximizing model in which preferences depend on health, one can derive an expression for the reduced form for health. Depending on the form for the utility function as well as constraints on health technology, etc., these final-form expressions for health are expressed as explicit or implicit functions of the predetermined variables. If we define mortality as the state of one's health dropping irreversibly below some critical value that is observed by the econometrician, then once we posit a distribution for the critical value we have a rule that links the conditional reduced-form expression for health to the probability of dying. The probability of dying in a finite period is just the hazard of dying in that period since it is conditional on having survived up to that time. Such considerations tie our analysis below to the general approach of standard demand analysis of health outcomes [e.g., Grossman (1972), Rosen and Taubman (1982), Rosenzweig and Schultz (1982a, b, 1983a, b), Behrman and Wolfe (1987), Wolfe and Behrman (1987)]. Given our particular data set and interests, however, the relations that we estimate are best viewed as conditional health demand functions in the sense of Pollak (1969, 1970), and not pure one-period reduced forms with only exogenous variables on the right side. That is, we are exploring the determinants of mortality hazards, conditional on a set of predetermined prior outcomes that are fixed throughout the sample period, such as education, longest occupation, or whose levels are determined by factors prior to the sample period such as pension income and expected Social Security income, or which are determined by factors that may change during the sample period but are assumed to be predetermined to the mortality hazard such as number of dependent children and the eligibility for, and hence expected dollar income of, supplemental income in 1975 and 1977 based on 1969 family eligibility criteria.

This specification of conditional death hazard functions raises the issue of simultaneity. Simultaneity bias may be a problem with all of these right-side variables, including education, because there may be persistent unobserved heterogeneity (e.g., genetic or family-background environment-related characteristics associated with inherent robustness, ability, and motivation) that affect outcomes throughout one's life. Several studies are consistent with the possibilities of such unobserved factors having influence at different points in the life cycle [e.g., Behrman, Hrubec, Taubman, and Wales (1980), Olneck (1977), Behrman and Wolfe (1984, 1987, 1989), Rosenzweig and Schultz (1983b), Wolfe and Behrman (1987)].

Within the health demand literature, the most common tradition is to emphasize the possible simultaneity bias for labor income. But the same possibility exists for nonlabor income (particularly if brighter people have greater labor market earnings and better investment strategies). To control for all such simultaneity with most such data sets, including the RHS, is difficult (and almost never done). The methods that usually are used to control for simultaneity, moreover, are not without their limitations since ideal instruments, which are highly correlated with the endogenous variable but orthogonal to the disturbance, are rarely available. If the former condition is not satisfied, measurement error bias may dominate in the estimates. If the orthogonality condition is not satisfied, simultaneity bias still may be a problem.

In this paper, we focus in our estimation upon sophisticated mortality analyses including the control for unobserved individual heterogeneity. Because of the lack of suitable instruments and the problems noted in the previous paragraph, we do not in addition worry about possible simultaneity biases. However, we do undertake and report on some estimates in which we drop or add some variable to see if that changes the estimated coefficients of other variables. Of course, even if simultaneity biases are important in our estimates, the associations that we uncover, while being biased indicators of causality, still can provide a good basis for predictions.

4. Statistical methods

Consider the continuous time duration model in which a nonnegative random variable T, time until death, has a density $f(t)$ and a cumulative distribution $F(t)$, both absolutely continuous. The hazard for T is the conditional density of T given $T > t \geq 0$ and is given by

$$\lambda(t) = f(t|T>t) = f(t)/[1-F(t)] \geq 0. \tag{1}$$

In terms of the integrated hazard, the density and distribution of T are

$$f(t) = \lambda(t)\exp\left\{-\int_0^t \lambda(\tau)\,\mathrm{d}\tau\right\} \tag{2}$$

and

$$F(t) = 1 - \exp\left\{-\int_0^t \lambda(\tau)\,\mathrm{d}\tau\right\}. \tag{3}$$

Let $\delta = 1$ if the duration is right-censored and $\delta = 0$ otherwise. The Dirac censoring distribution associated with realizations on δ is assumed to be independent of the survival time and is functionally independent of the survival distribution. The log-likelihood function is

$$\ln L = \sum_i f(t)(1-\delta) + \sum_i [1 - F(t)](\delta). \tag{4}$$

In the analysis to follow, we condition on a number of time-varying covariates and frailty differences among individuals that are not directly observable. Following Heckman and Singer (1984) and Manton et al. (1986), the conditional hazard is defined as

$$\lambda(t|x(t), \theta(t)) = \lim_{\Delta \to 0} \operatorname{Prob}[t < T < (t+\Delta)|T > t, x(t), \theta(t)]/\Delta, \tag{5}$$

where $x(t)$ are time-dated regressors that are assumed to be ancillary to T [Cox and Hinkley (1974)] and where $\theta(t)$ has density $\mu(\theta)$. The conditional duration density and the conditional duration distribution are constructed as above. In particular, we have

$$F(t|x, \theta) = 1 - \exp\left\{ -\int_0^t \lambda(\tau|x(\tau), \theta(\tau))\, \mathrm{d}\tau \right\} \tag{6}$$

and

$$f(t|x, \theta) = \lambda(t|x(t), \theta(t))(1 - F(t|x, \theta)). \tag{7}$$

The log-likelihood function is modified accordingly, and the Dirac censoring distribution is further restricted to be independent of θ.

Failure to control for unobserved frailities causes a downward bias in duration dependence. Moreover, misspecifying either the hazard or the frailty distribution leads to inconsistent estimates of the covariate effects. This is not a new point. The hazard process being modelled is highly nonlinear and a failure to properly specify the nonlinearity biases coefficient estimates [White (1980)]. Either ignoring or improperly specifying the distribution of measurement error in even linear models causes parameter estimates to be inconsistent. It is no surprise that the potential for both problems might put applied researchers in a very uncomfortable position when evaluating results using standard parametric estimators. The attention given to these attendent problems has focused on two separate approaches. The first, used by Manton et al. (1986), assumes a flexible parametric distribution for frailty differences among individuals that enters the hazard multiplicatively. The second, proposed by Heckman and Singer (1984), allows for the distribution of frailty differences to be estimated by a finite support general probability estimator [Keifer and Wolfowitz (1956)]. The estimator is consistent and approximate standard errors also can be generated [Heckman and Singer (1984)].

In the work that follows, we use a number of different parametric and semiparametric estimators of the hazard based on differing assumptions about the baseline hazard and survival distributions and about the distribution of frailty. We also examine the Maximum Penalized Likelihood Estimator (MPLE) as an alternative to estimators proposed by Manton et al. and Heckman and Singer. The appeal of the MPLE is that it reduces the impact of distributional assumptions about frailty and at the same time generates estimates that have a well-defined asymptotic normal distribution.

We consider both accelerated and proportional hazard models in our empirical work. The accelerated hazard models express the (natural) log of the date of death as a linear function of the covariates and of heterogeneity, $\log T = x(t)\beta + \theta(t) + \sigma\varepsilon$, where ε is a random disturbance and σ is a scale parameter. Failure time is $T = \exp\{x(t)\beta + \theta(t)\}T_0^{\sigma}$, where T_0 is an event time drawn from a baseline duration distribution for which the covariates and heterogeneity are zero. We consider two different distributions for T_0 that lead to two different expressions for the conditional hazard $(\lambda(t)|x(t), \theta(t))$. The first is the Weibull, which is the only baseline distribution that leads to a proportional hazard. The second is the log-logistic, which offers increased flexibility in the shape of the hazard, but at the cost of blurring somewhat the comparison of covariate effects with the proportional hazard estimates. The conditional hazards for the two baseline distributions are

$$\lambda(t|x(t), \theta(t)) = \sigma^{-1} \exp\{-x(t)\beta/\sigma - \theta(t)\sigma\} t^{(\sigma^{-1}-1)} \tag{8}$$

and

$$\lambda(t)|x(t), \theta(t) = \frac{\sigma^{-1} \exp\{-x(t)\beta/\sigma - \theta(t)/\sigma\}}{1 + \exp\{-x(t)\beta/\sigma - \theta(t)/\sigma\} t^{1/\sigma}}. \tag{9}$$

We consider two special cases of the Box-Cox conditional hazard utilized by Heckman and Singer for the proportional hazard specifications as well as two different ways in which frailty differences affect the hazard. The general forms for the conditional hazards are

$$\lambda(t|x(t), \theta(t)) = \exp\left\{x(t)\beta + \gamma\frac{t^k - 1}{k} + \theta(t)\right\} \tag{10}$$

and

$$\lambda(t|x(t),\theta(t) = \theta(t)\exp\left\{x(t)\beta + \gamma\frac{t^k - 1}{k}\right\}, \tag{11}$$

where, for $k = 1$, (10) and (11) reduce to Gompertz hazards in which time enters the log-hazard linearly, and where for $k = \infty$, time enters the log-hazards proportionally. The different specifications allow frailty differences to affect the log-hazard linearly (Heckman and Singer) or proportionally (Manton et al.).

We consider several different parametric distributions for $\theta(t)$ as well as semiparametric estimators for which the distribution of $\theta(t)$ need not be specified *a priori*, but it is either factored out of the likelihood function (Cox partial likelihood) or is estimated by a finite support density estimator. The latter estimators have been explored by Heckman and Singer. The parametric distributions that we use for $\theta(t)$ are the normal and the inverse Gaussian. We use the former for the first specification of the conditional hazard above and the latter for the second specification. As noted by Manton et al. (1986, p. 637), the inverse Gaussian provides a mixture that is quite flexible and allows for a very general description of the continuous variability in biological risks.

In order to facilitate direct comparisons of covariate effects among the different specifications, we normalize the factor loading for the finite support estimator at unity and utilize a standard normal mixture when using the Box–Cox conditional hazard. The inverse Gaussian is parameterized as

$$\mu(\theta(t)) = \left[\xi/\left(2\nu\pi\theta(t)^3\right]^{-1/2}\exp\left\{-\xi(\theta(t) - \nu)^2/(2\nu^2\theta(t))\right\}, \tag{12}$$

where we normalize the mean ν at unity and note that the parameter ξ is the reciprocal of the measure of dispersion. Since in Manton et al. frailty effects shift the hazard multiplicatively as with Cox's model, there is no need to normalize the factor loading at one in order to keep covariate effects in the same scale as with models in which no heterogeneity is considered explicitly. We also introduce an alternative method to handle the conditioning of heterogeneity in our estimation of reduced-form hazards that avoids the need to parameterize the mixing distribution, is consistent, and has a well-defined conditional limiting normal distribution. Since the estimator has not appeared widely in the econometrics literature, we now outline its form and motivate its appeal.

Maximum Penalized Likelihood Estimation (MPLE) avoids the problem of over-parameterization of heterogeneity, as well as some computational problems with alternatives such as the finite support generalized probability estimator employed by Heckman and Singer. MPLE is maximum likelihood with heterogeneity controlled by smoothing local variability in the density that has not been controlled by the covariates. The basic idea in MPLE is to smooth heterogeneity from the likelihood by including penalty terms that take into account the degree of roughness or local variability in the joint density of the data. The general form of a penalized log-likelihood under random sampling is given by

$$L_\alpha(f) = \sum_{i=1}^{n} \log f(x_i) - \alpha R(f), \tag{13}$$

where $f(x)$ is an unknown density, $\int f(x)\,\mathrm{d}x = 1$, $f(x) \geq 0$ for all x, $R(f) < \infty$, R is a functional, and α is the smoothing parameter. The choice of the smoothing parameter controls the balance between smoothness and goodness-of-fit, while the choice of the penalty functional, R, determines the type of behavior in the density estimate considered undesirable. For example, if R uses the first derivative, then R smoothes the slope of f. If R uses the second derivative, the curvature is smoothed as well. MPLE is a versatile method for our purpose because the functional form of R can be chosen according to various assumptions about the covariance structure of unobserved heterogeneity whose distribution is unknown. MPLE has a Bayesian interpretation since the choice of the smoothing parameter, α, determines the prior density on unobserved heterogeneity that is proportional to $\exp\{-R(f)\}$. Typically α is chosen by cross-validation methods. Maximum penalized likelihood then estimates the posterior density of f. With $\theta(t)$ affecting the log-hazard linearly, the penalized log-likelihood of the Box–Cox conditional hazard for single transition welfare spells can be written as

$$L_\alpha(f) = \sum_{i=1}^{n} (1-\delta_i)\log f_i(x_i) + \delta_i \log(1 - F_i(x_i)) - \|f(x)\|^2. \tag{14}$$

The penalty terms are written as

$$\|f(x)\|^2 = \sum_{j=1}^{s} \alpha \sum_{i=1}^{n} \{f^{(j)}(x_i)\}^2, \qquad \alpha > 0, \quad \text{for } j = 1,\ldots,s,$$

where

$$f(x) = \exp\left\{x(t)\beta + \gamma\frac{t^k - 1}{k} + \theta(t)\right\}$$

$$\times\left\{1 - \exp\left(-\int \exp\left[x(t)\beta + \gamma\frac{t^k - 1}{k} + \theta(t)\right]\mathrm{d}t\right)\right\},$$

where $\delta_i = 1$ if ith individual is censored and 0 if uncensored, j denotes the degree of the derivative, α_j is a corresponding smoothing parameter, and $f(x)$ is the joint probability density of failure times and heterogeneity.

Elsewhere [Huh and Sickles (1989)] the existence of a unique maximum for the MPLE has been proven in the Sobolev space. The proofs are based on the property that if the set of points of support for the heterogeneity distribution is a subset of the Hilbert space, then the Hilbert space defined on any arbitrary interval (a, b) is a Reproducing Kernel Hilbert Space for all values of the mixing random variable in that interval. Since the space of intervals is a closed-convex subset of a Hilbert space, and since (14) is continuous, the second Gateaux variation of (14) is uniformally negative definite and thus a maximizer of (14) exists. Huh and Sickles also have proven existence if heterogeneity is correlated with the covariates and if heterogeneity is both individual- and time-specific.

MPLE would appear to have an advantage over NPMLE when the censoring rate is relatively high and the distribution of heterogeneity is long-tailed. The reason is that the mass point method in practice appears unable to identify such distributions with the small number of points of support necessary to implement the estimator. MPLE, however, is designed to identify such roughness and smooth it using spline functions, thus removing a potential source of instability in estimation.

5. Results

We first consider the hazard data for black males and for white males, smoothed versions of which are given in fig. 1.[3] (See also table 4.) It is evident that the hazard is higher for blacks at every age covered. Thus, in our limited age range, there is no cross-over in hazard rates for blacks and whites.

Our estimated conditional demand relations are presented in tables 1–3. The first two tables give estimates of the effects of changes in the covariates

[3]The figure is limited to people aged 60 through 66 (even though, as we note in section 2, our data include ages 58–73) in part because of the smaller sample sizes for other ages (arising from the age and panel structure of the RHS) and in part because of the incomplete information on death after 1977.

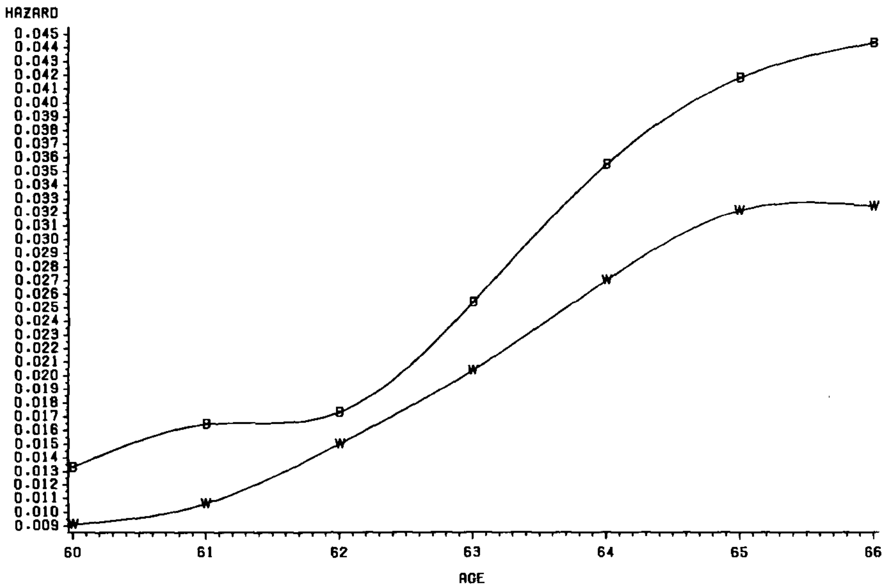

Fig. 1. Hazard functions for blacks (B) and nonblacks (W).

on the log-hazard for blacks and whites. We have estimates with and without allowance for heterogeneity. Numbers in parentheses are asymptotic t-statistics.

Table 1 is for black men. In columns 1 and 2, we present the results assuming a Weibull and log-logistic accelerated time to failure model. The estimates indicate the associations of the right-side variables, evaluated at their mean, with the log-hazard, which yields comparability with the estimates in the various proportional hazard models. The statistically significant variables are for marital status and pension income. The marital status estimates indicate that those who are married or divorced/separated have significantly less probability of dying in a given age range than those who are never married or who are widowed. The pension effect presumably indicates the advantage of higher income or related characteristics rather than occupation per se since longest occupation is included as two dichotomous variables for professional and management. However, Social Security benefits, Supplemental Security income,[4] education, occupation, and number of children are not statistically significant. As noted in section 3, all of these coefficient estimates may be affected by simultaneity. Marital status, for example, may

[4]This variable may have measurement error since some people died before the program began in 1974 and since it is age-related.

Table 1

Hazard estimates for black men; sample size = 692; t-statistics in parentheses.

Variable	Weibull accelerated hazard	Log-logistic accelerated hazard	Cox[a] proportional hazard	Weibull proportional hazard (NPMLE)	Weibull proportional hazard normal frailty	Weibull proportional hazard (MPLE)	Gompertz proportional hazard inverse Gaussian frailty
Intercept	−0.764 (0.85)	−0.277 (0.41)	−0.456 (6.11)	−0.906 (0.20)	−0.383 (1.21)	−0.840 (0.92)	−1.54 (2.83)
Education	−0.0727 (1.08)	−0.0836 (1.20)	−0.068 (1.06)	−0.079 (1.16)	−0.096 (1.17)	−0.138 (1.21)	−0.0711 (1.11)
Education^2	0.002 (0.47)	0.0023 (0.51)	0.003 (0.60)	0.002 (0.46)	0.003 (0.50)	0.0102 (0.71)	0.00181 (0.43)
Married	−2.630 (10.7)	−2.725 (11.3)	−2.536 (10.7)	−2.796 (5.32)	−3.191 (11.2)	−2.777 (11.4)	−2.622 (11.05)
Widowed	0.016 (0.03)	−0.0517 (0.12)	0.053 (0.31)	0.009 (0.05)	−0.007 (0.03)	0.0333 (0.52)	0.00994 (0.06)
Divorced/separated	−2.556 (4.84)	−2.68 (4.91)	−2.491 (4.85)	−2.721 (3.85)	−3.128 (5.33)	−2.718 (4.66)	−2.55 (4.97)
Longest occupation professional	0.356 (1.47)	0.287 (1.25)	0.263 (1.02)	0.379 (1.32)	0.400 (1.16)	0.337 (1.36)	0.344 (1.32)
Longest occupation management	0.671 (1.30)	0.850 (1.65)	0.585 (1.48)	0.702 (1.35)	0.765 (1.22)	0.705 (1.44)	0.650 (1.63)

Table 1 (continued)

Variable	Weibull accelerated hazard	Log-logistic accelerated hazard	Cox[a] proportional hazard	Weibull proportional hazard (NPMLE)	Weibull proportional hazard normal frailty	Weibull proportional hazard (MPLE)	Gompertz proportional hazard inverse Gaussian frailty
Pension income (1000$)	−0.331 (2.50)	−0.330 (2.58)	−0.330 (2.53)	−0.335 (2.66)	−0.394 (3.12)	−0.348 (2.61)	−0.327 (2.43)
Expected Social Security benefits (1000$)	0.020 (0.31)	0.0714 (0.28)	−0.008 (0.13)	0.032 (0.60)	0.0492 (0.81)	0.0610 (0.91)	0.0455 (0.78)
Number of dependet children	0.028 (0.42)	−0.0216 (0.35)	0.009 (0.15)	0.012 (0.12)	−0.005 (0.04)	0.0582 (0.22)	0.0250 (0.39)
Supplemental Security income (1000$)	−0.798 (1.35)	−0.785 (1.38)	−0.743 (1.28)	−0.797 (0.99)	−0.803 (0.97)	−0.797 (1.38)	−0.758 (1.29)
ln duration	0.559 (5.07)	0.511 (4.60)	0.436 (4.56)	0.385 (1.32)	0.587 (4.51)	0.520 (4.11)	0.402 (2.97)
ln L	−399.40	−389.98	−400.4	−395.83	−392.43	−397.2	−404.8
χ^2	286.2	305.4	287.2	293.7	300.6	296.6	275.4

[a]See footnote 5.

be correlated with the unobserved health stock with ill men being less likely to marry to remarry even though we use marital status as of 1969 while most of the deaths occur much later. Therefore, we have reestimated the equations and included a self-assessed measure of health as of 1968, a variable that is highly correlated with physician evaluations in this age range. The coefficients, including those for marital status, are nearly the same as those shown in table 1.

Columns 3–7 present various proportional hazard estimates that allow for nonparametric and normal heterogeneity in the Weibull model (columns 4 and 5) and inverse Gaussian frailty (column 7). The numerical results are very similar to those in columns 1 and 2. Coefficient estimates are very robust though significance levels change and the duration estimate is insignificant using the Nonparametric Maximum Likelihood Estimator (NPMLE). The estimates with no heterogeneity and a log-logistic accelerated hazard fit the data best in terms of the maximized value of the log-likelihood function. We again find strong negative associations with the hazard of pension income and of being married or divorced/separated. These associations again persist even if we control for self-assessed health status (in results not shown).

The results for the divorced group are surprising given earlier studies, using different statistical methods, such as Rosen and Taubman (1982) that indicate that married men are healthier than divorced men. A possible interpretation is that single men never marry because their poor health makes them poor potential marriage partners and that widowers do not do as well as those who are divorced or separated because of their grief or because there is assortative mating on unmeasured characteristics that affect morbidity. However, this explanation is at odds with the finding that the inclusion of self-assessed prior health has little effect on the marital status results. An alternative story for married and divorced/separated men is that one's current and past wife improved one's health *stock* via reducing poor habits, nagging husbands to see a doctor, and nursing. In this interpretation the better stock of health has not yet depreciated for the divorced/separated men, though the trauma of one's spouse's death takes a toll on those widowed.

Table 2 contains the corresponding results for white males. The greater significance levels than in table 1 partly reflect the approximate tenfold increase in sample size. Coefficient estimates again differ little across the models. As was found with the black men, introduction of nonparametric or parametric heterogeneity yields a small improvement in fit, similar parameter estimates, and changed significance levels (smaller in column 4). For the estimates for white men, all three marital categories have highly significant coefficient estimates with widows having an increased hazard relative to being never married. Increased pension income significantly reduces the hazard. Other variables tend to have expected signs, but are not statistically

Table 2

Hazard estimates for white men; sample size = 692; t-statistics in parentheses.

Variable	Weibull accelerated hazard	Log-logistic accelerated hazard	Cox[a] proportional hazard	Weibull proportional hazard (NPMLE)	Weibull proportional hazard normal frailty	Weibull proportional hazard (MPLE)	Gompertz proportional hazard inverse Gaussian frailty
Intercept	−2.129 (8.74)	−1.63 (6.59)	−1.79 (64.5)	−2.366 (2.19)	−1.979 (10.7)	−1.88 (7.15)	−2.757 (15.1)
Education	−0.026 (0.85)	−0.004 (0.16)	0.045 (1.36)	0.015 (0.47)	0.004 (0.10)	0.0221 (0.79)	−0.0193 (0.75)
Education^2	−0.0012 (0.85)	0.00031 (0.34)	−0.002 (0.98)	0.0004 (0.23)	0.0001 (0.06)	0.000031 (0.45)	0.0009 (0.55)
Married	−1.738 (23.0)	−1.740 (22.6)	−1.687 (21.8)	−1.796 (16.3)	−2.005 (22.7)	−1.872 (20.7)	−1.955 (19.9)
Widowed	1.228 (22.4)	1.333 (24.0)	1.172 (20.0)	1.350 (16.6)	1.530 (21.1)	1.441 (19.3)	1.395 (20.2)
Divorced/separated	−1.241 (5.20)	−1.331 (5.47)	−1.284 (5.35)	−1.297 (4.94)	−1.499 (5.45)	−1.531 (5.21)	−1.426 (5.05)
Longest occupation professional	−0.098 (1.32)	−0.131 (1.66)	−0.119 (1.56)	−0.107 (1.34)	−0.132 (1.45)	−0.122 (1.47)	−0.123 (1.52)

Longest occupation management	−0.104 (1.49)	−0.141 (1.92)	−0.130 (1.77)	−0.117 (1.53)	−0.153 (1.72)	−0.151 (1.62)	−0.129 (1.77)
Pension income (1000$)	−0.186 (7.10)	−0.184 (7.12)	−0.179 (7.01)	−0.186 (8.42)	−0.205 (8.6)	−0.199 (8.19)	−0.201 (7.81)
Expected Social Security benefits (1000$)	0.028 (1.59)	0.0295 (1.62)	−0.016 (0.82)	0.030 (1.64)	0.0303 (1.50)	0.0281 (1.49)	0.0533 (1.63)
Number of dependent children	−0.024 (0.42)	−0.059 (0.93)	0.040 (0.72)	−0.029 (0.46)	−0.046 (0.67)	−0.0441 (0.98)	−0.0488 (0.88)
Supplemental Security income (1000$)	0.099 (0.48)	0.466 (0.38)	0.174 (0.84)	−0.001 (0.003)	−0.037 (0.16)	−0.0227 (0.92)	−0.0020 (0.12)
ln duration	0.552 (15.6)	0.543 (14.9)	0.435 (8.72)	0.359 (5.36)	0.532 (12.5)	0.521 (11.8)	0.413 (13.2)
ln L	−3815.5	−3772.5	3820.4	−3814.4	−3808.0	−3813.2	3836.1
χ^2	2214.1	2300.1	2219.0	2216.3	2229.1	2218.8	2173.1

[a] See footnote 5.

Table 3

Differences in racial means of selected variables and impacts on the hazard.

Variable	Black–white means	Percentage effect on log hazard[a]
Married in 1969	−0.0597	10.1
Widowed in 1969	0.0296	3.46
Divorced/separated in 1969	0.0571	−7.33
Education	−3.24	−14.6
(Education)^2	−52.5	10.5
Longest occupation professional	−0.0756	0.90
Longest occupation management	−0.151	1.96
Expected Social Security benefits in 1973	−0.279	0.5
Pension income	−0.442	7.92
Dependent children in 1973	0.256	1.02
Supplemental Security income in 1975	0.188	3.27

[a]Using Cox partial likelihood estimates for whites in table 2.

significant at conventional levels. A comparison of tables 1 and 2 indicates that the coefficients significant in both tables are usually larger in absolute value for blacks than for whites though the coefficients for being widowed are larger for whites than for blacks (and only significant for the former). The duration estimates, however, are similar for the two groups.

We have examined several specifications (not presented) in which data from the two racial groups were pooled. These models were estimated using the Cox partial likelihood, the Weibull accelerated hazard, and the NPMLE model. We examined models in which selected regressors, including the constant term, were allowed to differ between the two groups, although we did not consider a model in which race interacted with all covariates due to computational constraints. The coefficients typically lay between the estimates for blacks and whites. In all pooled models, however, we rejected the null hypothesis that the coefficients were the same for blacks and whites at the 99 percent level.

Table 3 uses the white Cox partial likelihood estimates from table 2 to assess the association of the racial differences in the means with the hazard rate differentials. We use the white rather than the black hazard because the former is more precisely estimated with its ten times larger sample. The big differences in table 3 come from marital status, pension income, and education (whose coefficients are not statistically significant and whose linear and square terms largely offset each other). Overall the white hazard would be about 19 percent higher if whites had the blacks' observed characteristics. Approximately the same results would be found in the other proportional hazards given the robustness of the coefficients.[5]

[5]However, the white hazard would be about 11 percent higher if whites had blacks' observed characteristics rather than the 18 percent figure when differences in pension income are based on the 1975 figures. See note 2 above.

Table 4
Black–white hazard rates.

Age	White hazard rate	Black hazard rate	Percentage difference
60	0.0091	0.0133	46.2
61	0.0106	0.0164	54.7
62	0.0150	0.0173	15.3
63	0.0204	0.0254	24.5
64	0.0270	0.0355	31.5
65	0.0321	0.0418	30.2
66	0.0324	0.0443	36.7

The actual differences in the hazards are given in table 4 for ages 60–66. The differentials range from 35 to 15 percent with some instability arising from small subsamples, especially for blacks. The average differential is about 34 percent. The differential in the median time period is about 25 percent. Differences in the observed characteristics are associated with between 60 and 80 percent of the difference in the hazard rates.

6. Conclusion

In this paper we have explored inequalities in mortality between black and white older adult males in the United States. We have estimated hazard functions separately for blacks and whites. The equations have different coefficients by race. Within a race, the equations are robust to changes in specification including allowance for heterogeneity. Replacing whites' means by blacks' means in the proportional hazard for whites would raise the white hazard rate by about 19 percent, a noticeable amount that is consistent with most of the inequalities in the observed mortality hazards. Such observed characteristics – particularly those related to marital status, pension income, and education – thus capture most of the black–white mortality differences among older men. The factors emphasized by Jaynes and Williams (1989, p. 425) in the quotation given in the introduction – including poorer quality health care, greater exposure to environmental risk factors and the stress of prejudice and discrimination – if important, apparently largely work through these observed characteristics. If there were movements towards convergence in regard to such observed socioeconomic characteristics and their covariates, therefore, there probably would be a reduction in older adult male black–white mortality rate inequalities.

References

Andrisani, Paul J., 1977, Internal–external attitudes, personal initiative and the labor market experience of black and white men, Journal of Human Resources 12, 308–328.

Ashenfelter, Orley, 1977, Comment on Smith–Welch, black/white male earnings and employment: 1960–1970, in: Thomas Juster, ed., The distribution of economic well-being (Ballinger, Cambridge, MA) 296–298.

Behrman, Jere R., Zdenak Hrubec, Paul Taubman, and Terance Wales, 1980, Socioeconomic success: A study of the effects of genetic endowments, family environment and schooling (North-Holland, Amsterdam).

Behrman, Jere R., Robin C. Sickles, and Paul Taubman, 1988, Age-specific death rates, in: Rita Ricardo-Campbell and Edward P. Lazear, eds., Issues in contemporary retirement (Hoover Institution, Stanford, CA) 162–190.

Behrman, Jere R. and Barbara L. Wolfe, 1984, The socioeconomic impact of schooling in a developing country, Review of Economics and Statistics 66, 296–303.

Behrman, Jere R. and Barbara L. Wolfe, 1987, How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation?, Journal of Econometrics 36, 185–204.

Behrman, Jere R. and Barbara L. Wolfe, 1989, Does more schooling make women better nourished and healthier? Adult sibling random and fixed effects estimates for Nicaragua, Journal of Human Resources 24, 644–663.

Cooper, R., M. Steinhauer, A. Schatzkin, and W. Miller, 1981, Improved mortality among U.S. blacks, 1968–1978: The role of antiracist struggle, International Journal of Health Services 11, 511–522.

Cox, D. and D. Hinkley, 1974, Theoretical statistics (Chapman and Hall, London).

Darity, William A., 1982, The human capital approach to black–white earnings inequality: Some unsettled questions, Journal of Human Resources 18, 72–93.

Duleep, Harriet, 1986, Measuring income's effect on adult mortality, Journal of Human Resources 21, 238–251.

Freeman, Richard B., 1973, Changes in the labor market for black Americans 1948–1972, Brookings Papers on Economic Activity 1, 67–120.

Good, I.J. and R.A. Gaskins, 1971, Nonparametric roughness penalties for probability densities, Biometrica 58, 255–277.

Grossman, Michael, 1972, The demand for health: A theoretical and empirical analysis (Columbia University Press, New York, NY).

Heckman, James and B. Singer, 1984, A method for minimizing the impact of distributional assumptions in econometric models for duration data, Econometrica 52, 271–320.

Huh, K. and Robin C. Sickles, 1989, Estimation of the duration model by maximum penalized likelihood and simulation based estimation, Mimeo. (Rice University, Houston, TX).

Jaynes, Gerald David and Robin M. Williams, Jr., 1989, A common destiny: Blacks and American society (National Academy Press, Washington, DC).

Kahn, Lawrence M. and Peter D. Sherer, 1988, Racial differences in professional basketball players compensation, Journal of Labor Economics 6, 40–61.

Kiefer, J. and J. Wolfowitz, 1956, Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters, American Mathematical Statistics 27, 887–906.

Kitigawa, E. and P. Hauser, 1973, Differential mortality in the United States of America: A study of socioeconomic epidemiology (Harvard University Press, Cambridge, MA).

Maasoumi, Esfandiar, 1986, The measurement and decomposition of multi-dimensional inequality, Econometrica 54, 991–997.

Maasoumi, Esfandiar and G. Nickelsburg, 1988, Multivariate measures of well being and an analysis of inequality in the Michigan data, Journal of Business and Economic Statistics 6, 327–334.

Manton, K.G., E. Stallard, and J.W. Vaupel, 1986, Alternative models for the heterogeneity of mortality risks among the aged, Journal of the American Statistical Association 81, 635–644.

Olneck, Michael, 1977, On the use of sibling data to estimate the effects of family background, cognitive skills, and schooling: Results for the Kalamazoo brothers study, in: Paul Taubman, ed., Kinometrics: Determinants of socioeconomic success within and between families (North-Holland, Amsterdam) 125–162.

Orazem, Peter T, 1987, Black–white difference in schooling investment and human capital production in segregated schools, American Economic Review 77, 714–723.

Pollak, Robert A., 1969, Conditional demand functions and consumption theory, Quarterly Journal of Economics 83, 60–78.

Pollak, Robert A., 1970, Habit formation and dynamic demand functions, Journal of Political Economy 78, 745–763.

Rosen, Sherwin and Paul Taubman, 1982, Healthiness, education and marital status, in: Victor Fuchs, ed., Economic aspects of health (University of Chicago Press, Chicago, IL) 121–142.

Rosenzweig, Mark R. and T. Paul Schultz, 1982a, The behavior of mothers as inputs to child health: The determinants of birth weight, gestation and rate of fetal growth, in: V. Fuchs, ed., Economic aspects of health (University of Chicago Press, IL) 53–92.

Rosenzweig, Mark R. and T. Paul Schultz, 1982b, Child mortality and fertility in Colombia: Individual and community effects, Health policy and education 2 (Elsevier Scientific Publishing Company, New York, NY).

Rosenzweig, Mark R. and T. Paul Schultz, 1983a, Consumer demand and household production: The relationship between fertility and child mortality, American Economic Review 73, 38–42.

Rosenzweig, Mark R. and T. Paul Schultz, 1983b, Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight, Journal of Political Economy 91, 723–746.

Shulman, Steven, 1987, Discrimination, human capital, and black/white unemployment, Journal of Human Resources 22, 361–376.

Sickles, Robin C. and Paul Taubman, 1986, An analysis of the health and retirement status of the elderly, Econometrica 54, 1339–1356.

Smith, James P., 1984, Race and human capital, American Economic Review 74, 685–698.

Smith, James P. and Finis Welch, 1977, Black-white male wage ratios: 1960–1970, American Economic Review 67, 323–338.

Welch, Finis, 1973, Black-white differences in returns to schooling, American Economic Review 63, 893–907.

Welch, Finis, 1974, Education and racial discrimination, in: Orley Ashenfelter and Albert Rees, eds., Discrimination in labor markets (Princeton University Press, Princeton, NJ).

White, Halbert, 1980, Using least squares to approximate unknown regression functions, International Economic Review 21, 149–170.

Wolfe, Barbara L. and Jere R. Behrman, 1987, Women's schooling and children's health: Are the effects robust with adult sibling control for the women's childhood background?, Journal of Health Economics 6, 239–254.

Yi et al., 1986, CTM: A user's guide, NORC manuscript.