ISSUES IN CONTEMPORARY RETIREMENT

Rita Ricardo-Campbell
Edward P. Lazear
Editors

HOOVER INSTITUTION PRESS
STANFORD UNIVERSITY STANFORD, CALIFORNIA

Contents

PART III: POLICY IMPLICATIONS

6

Age-Specific Death Rates

Jere R. Behrman, Robin Sickles, and Paul Taubman

Good knowledge of why variations in mortality occur is important. Such variations might underlie behavioral differences in labor force, retirement, health, medical care use, and saving patterns; they might also influence public policy decisions such as what transfer payments should be made and to whom these payments should be made.

Economists lately have investigated in detail the determinants of age of retirement.[1] Economists and other social scientists also have analyzed the determinants of mortality.[2] The mortality studies are generally based on a single cross-section or a short panel. In these studies, the probability of dying in a given time period is related to variables such as the person's education and age.

Recently, however, demographers and statisticians have developed and improved alternative statistical models to estimate baseline survivor distributions (that is, what percentage of the sample is alive at a given age), which can shift with covariates over time.[3] A major problem, which these methods are designed to overcome, is that in most samples not everyone has died by

This work has been supported by grants from the National Institute on Aging and from the Hoover Institution. The authors would like to thank Abdo Yazbeck for his valuable computational assistance and Jennifer Green for her editorial help. Special thanks go to Richard Burkhauser for his many suggestions.

the end of the sample frame. This right censoring is generally dealt with by making distributional assumptions about the form of the survivor function, using an unobserved baseline and employing a partial likelihood estimator, or using nonparametric techniques. Another major problem is the potential bias caused by unobserved heterogeneity in individual frailties that appears less severe as the number of explanatory variables is increased.

Previous related studies by demographers have generally been based on small samples and short periods of time. In this paper we make use of two large samples with data on death by month for ten and sixteen years, respectively. After presenting these data sets, we discuss the models we estimate and the construction of our variables, and we present results based on several estimators and different distributional assumptions.

The Data

The first of our two data sets is the Retirement History Survey (RHS), which has been used extensively by economists. The RHS is a large random sample with relatively rich information and with about ten years of death data. A large proportion of its observations are censored, which may make our results sensitive to the distributional assumptions used. We also consider a second data set, the Dorn sample, which is unknown to (or at least has not been used by) most economists. This is quite a large sample of white men who were generally born before 1900 and for whom censoring is less of a problem than for the RHS. However, with the exception of smoking and occupational data, less covariate information is available in the Dorn sample. Finally, the Dorn sample contains a nonrandom element that we discuss below.

The Retirement History Survey

The Retirement History Survey was started in 1969 with about 11,000 men and women. 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 (and shortly we will be able to extend coverage through 1979). Death information has been collected from two sources. The RHS records death as a reason for non-reinterview when they know this to be the case, generally through interviewing the widow. This source is incomplete. The other source is the Social Security files, which record death reported to them by month and year as part of the process of issuing death benefits (such as burial grants and survivor benefits for dependent children) and making necessary adjustments in Old Age, Survivor and Disability Insurance benefits. We currently have this death information through 1977 (with incomplete data into 1979).[4] We have compared these two sources, and there are only two cases of deaths recorded in the RHS that are not in the Social Security files. Moreover, the Social Security files' date of death is in accord with the RHS in that the individual does not give interviews after Social Security files. Moreover, the Social Security files' date of death is in accord with the RHS in that the individual does not give interviews after Social Security records his or her death. Duleep (1986), following up on the earlier work of Rosen and Taubman (1984), has indicated that in comparison to national death rates the Social Security files now record nearly all deaths.

The RHS contains substantial information on the respondents and their spouses, including age, education, wealth, current earnings, pensions, 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, medical usage, retirement status and plans, nutrition, and some aspects of life style including contact with children.

The advantages of the RHS include: (1) the RHS contains substantial information on respondents and spouses; (2) it is a random draw of the population of heads of households in 1969; and (3) it is of manageable size to experiment with various hazard functions employing different techniques and assumptions with a relatively large number of covariates. The major disadvantage of the RHS is that currently we can only look at the survivor curve over at most a fifteen-year age-time period and, when we wish to study each age group separately, a ten-year period. In addition, a large proportion (about 80 percent of the sample was still alive in 1977 and is thus censored in an analysis of survival. Below we report how sensitive our estimates are to alternative treatments of the censoring problem.

The Dorn Sample

The sample was originally constructed by Dorn (1958), whose pioneering effort was extended by Kahn (1966), by Rogot (1974)—who is our source for the following description—and by Rogot and Murray (1980); it has recently been updated through 1980 by Hrubec, Norman, and Rogot.[5] Dorn was interested in studying the relationship of tobacco use to mortality experience in general and to specific causes of death. With the cooperation of the Veterans Administration (V.A.), he mailed a short questionnaire to 293,958 U.S. veterans who in December 1953 held U.S. government life insurance policies and who had served in the armed forces between 1917 and 1940. The questionnaire, which can be found in Kahn (appendix E), asked how many times a day a person smoked cigarettes, cigars, and/or a pipe, how long ago he had stopped, his occupation, his industry, and his age. About 200,000 veterans responded in 1954. Another 49,000 responded to a second mailing in January 1957, and 46,000 did not reply. The V.A. recorded deaths by month as well as by year and by causes.

Rogot (1974, 192) presents the age distribution for 1954 and 1957 as:

Age	*Respondents in 1954*	*Respondents in 1957*	*No reply*	*Total*
30–34	7,421	43	2,148	9,612
35–44	16,735	7,156	4,037	27,928
45–54	10,317	1,242	2,232	13,791
55–64	137,820	26,579	31,468	195,867
65–74	25,002	13,683	5,603	44,288
75–84	1,525	523	424	2,472
30–84	198,820	49,226	45,912	293,958

People of different ages may have had different environmental exposures and different medical technology available when ill. Thus, it is possible that survivor functions vary by age cohort. This sample is large enough to allow the estimation of separate survivor functions for different age groups. Those in the 55–64 cohort can eventually be divided into single-year intervals. For other ages we can combine several adjacent groups. Currently we subdivide those born before 1891 from those born between 1891 and 1900.

We have excluded from the analysis those people who have no information on any of the variables studied. This left us with a sample of nearly 200,000. We have examined the plots of the age-specific death rates for the 200,000 and for the full sample. They are nearly identical.

The Dorn sample has cause-of-death data taken from death certificates. The cause-of-death data will be aggregated into heart disease, stroke, all forms of cancer, accidents, and all other. Although we realize that the information on cause of death is noisy, in part because of the difficulty in distinguishing immediate primary from underlying and contributory causes, the categories given are fairly broad and the first four categories include the causes of most deaths. We will eventually modify our hazard analysis to treat each cause as a distinct outcome, just as economists have treated employment, unemployment, and out of the labor force as separate states.[6]

The Dorn veteran sample includes only those veterans whose V.A. life insurance was still in force in December 1953. A questionnaire administered at that time collected information on occupation. As shown by Dorn (1958) and by Kahn (1966), the sample has disproportionally fewer unskilled workers than the corresponding white male cohort.[7]

During the last several decades, epidemiologists have investigated the accuracy of the V.A.'s information on date of death (see DeBakey and Beebe 1952; Beebe and Simon 1969; Cohen 1953). Procedural details differ across these studies, but basically the researchers took death certificates of men in the appropriate age range, matched them to military records to obtain military serial numbers, and then gave the names and numbers to the V.A. Roughly 95 percent of the deaths were recorded in the V.A. files, with many of those not listed having been dishonorably discharged or in the army no more than four days during World War I. This high rate of coverage occurs because veterans draw pension benefits that cease at death and other benefits that commence at death, such as burial plots, a flag, and a burial allowance. In the Dorn sample the economic incentives to keep in touch with the V.A. were particularly strong, since all participants had V.A. life insurance in force in 1954. Although some 75,000 terminated their insurance between 1963 and 1969, they were still eligible for the death benefits. For the Dorn sample, Rogot (1974, 190) reports that special efforts were made to check these 75,000 cases, and he says, "The overall mortality follow-up, with respect to the fact of death and year of death, is considered to be almost 100 percent complete."

The data on death have been updated periodically. The data for 1969–1980 have been collected and added to the file. However, we only received the post-1969 data in early 1987. Our current analysis is therefore based on death records through 1969. Moreover, we currently are limited by computer memory size to using about 85,628 individuals.

Sample means and standard deviations (S.D.) are given for the Dorn and RHS samples in Appendixes A and B, respectively.

Models and Variables

In this section we provide a brief overview of our models and variable construction. A detailed treatment of the statistical models is given in Appendix C.

We estimate both proportional and accelerated hazard models. The hazard rate is defined to be the probability of dying in a year divided by the probability of being alive at the beginning of the year. (More formally this is the ratio of the density function to one minus the cumulative density function.) In the proportional hazard model covariates act multiplicatively on the baseline hazard, whereas in the accelerated hazard model the covariates act multiplicatively on the baseline failure time. Although the potential for individual specific heterogeneity exists in both data sets, we do not deal with the problems by introducing flexible parametric distributions for the heterogeneity (Manton, Stallard, and Vaupel 1986) or by estimating the distribution function for heterogeneity (Heckman and Singer 1984a, 1984b). Rather, as a first attempt at examining the patterns of covariate effects in both data sets, we deal with the potential for omitted variable bias as one typically does in any nonlinear modeling environment—that is, by including a rather exhaustive set of individual specific explanatory variables, or covariates in the terminology of the demographers, and by experimenting with different functional forms. The proportional hazard estimates are based on maximizing the Cox partial likelihood. The accelerated hazard estimates are based on maximum likelihood using several flexible parametric baseline survivor distributions: exponential, Weibull, lognormal, loglogistic, and gamma. A more detailed discussion of the statistical models is given in Appendix C.

We should point out that, due to the construction of the Dorn sample, all sample covariate information is fixed. This is not the case in the RHS, and thus time-varying covariates are considered in our estimates. Furthermore, both samples are right censored, since not all individuals die by the end of the sample period. We do not, however, deal with the issue of left censoring. In both samples, respondents are not followed from birth, but rather are on average 57 years old when they enter the Dorn sample in 1954 and are no younger than 58 when they enter the RHS in 1969. Thus, inferences drawn on the effects of the covariates for individuals at all age groups would most likely be biased. Dealing with this potential shortcoming by truncating the densities of failure times (which truncated densities would be censored for those with incomplete spells) is a computational complication that we hope to deal with in future research.

The economic model that justifies inclusion of our covariates is quite well known and is outlined in Sickles and Taubman (1986). Although it is a static model, it does explain the role of the explanatory variables in the reduced form for the hazard function. An individual maximizes a utility function whose arguments are consumption, leisure, and health, subject to a budget constraint and a health production function. First-order conditions for utility maximization are then used to solve for the reduced form demand for consumption, hours of work, and health. Here we consider only health in isolation. The general specification of the health augmentation equation is based on the work of Anderson and Burkhauser (1983), Grossman (1972), and Taubman and Rosen (1982). The unobservable health stock is endogenously determined and can be augmented by investment in health services or depreciated by the environment of the home and work place. The health stock differs across individuals and families and is determined in part by social and demographic factors such as education, occupation, race, and age. Other important variables include access to health information and the ability to pay for health services as well as life style and genetic factors.

We use as an index of healthiness the hazard of dying. In earlier work with the RHS, the index of heathiness was self-reported health status relative to others the same age. Our proxies for the determinants of health are somewhat limited in the Dorn sample. Occupation is represented by an index for the riskiness of the work place. In the Dorn sample we have industry and occupation at the three-digit level. With the use of the *Underwriters' Handbook* used by the life insurance industry, we can replicate the risk class attached to these industry and occupation classifications when insurance premiums are set. The index ranges from 1 to 7 as the riskiness of the occupation rises and has a mean of 2.2 with some variation across age cohorts. The riskiest occupations are jobs such as firefighting and police work, whereas the least risky are jobs such as teaching. Another index was developed to control for the physical activity of the occupation. The index ranges from 1 for sedentary to 4 for heavy construction jobs. Life style is represented by information on tobacco use. The Dorn sample contains this information in the form of how much tobacco the respondent used, how many years he used it, and the manner in which he used it. At this stage we have constructed variables that indicate the number of years of occasional usage (an average of 3 percent of the years individuals were over ten years old) and the number of years of regular usage (an average of 54 percent of the years individuals were over ten years old). Since the questionnaire was administered in the mid-1950s when the respondent ages differed, it is possible that younger men who smoked regularly throughout their lives would still be able to report fewer years they smoked regularly than comparable older men. To account for this, we have divided the years-smoked variables by "age ten" at the survey data. These normalized variables have noticeably higher "t"-test statistics in otherwise identical equations. This adjustment may not be perfect; hence, the tables based on narrower age cohorts may be more appropriate for these variables. It is possible to distinguish type of tobacco use, but we have not yet done so. Nor have we made use of all information on past versus current habits. Factors such as regional differences in access to health care and exposure to health-depreciating environments are controlled for by four regional dummy variables with the West the omitted category.

There is a more extensive list of variables in the RHS that can be constructed to represent determinants of health than in the Dorn sample. Many of the variables we use are the same as those used in Sickles and Taubman (1986). We have socioeconomic data on race and marital status (we have omitted women from the study), number of years of education, longest occupation at the two-digit level, number of dependents, and genetic proxies such as the number of surviving parents of the respondent heads of households. We also have economic data such as initial wealth holdings, Social Security benefits, wage and nonwage income, income from assets, and Supplemental Security Income (SSI). A few words are in order about why we included some and excluded other economic variables, in particular earnings. Earnings, retirement, and health are intercorrelated, probably in a causal way.[8] Therefore, inclusion of earnings would probably lead to simultaneous equation estimation problems that we are not yet in a position to examine. To avoid this problem with Social Security benefits, we calculated the benefits that a person would expect to receive if retirement began in the respective year on the basis of earnings histories before 1969. They are computed using covered earnings taken from each person's Social Security record, which is part of the RHS, and then replicating Social Security's rules. By using benefits available rather than those actually paid to retirees, we avoid an obvious selection problem.

SSI, which began in 1974, was based on eligibility criteria as of 1969. The time-varying covariates in the analysis are number of dependents, number of surviving parents, potential Social Security benefits, and SSI. The remaining covariates are fixed throughout the sample period.

Results

We turn first to results based on the Dorn sample. Table 6.1 contains the maximum likelihood estimates of the accelerated hazard model for the period 1954–1970 using a sample of 85,628 persons. In these estimates the baseline hazard is the gamma function, which is often our best fitting model. We present the results in terms of elasticities of the time to death with respect to selected covariates.

About 65 percent of the Dorn sample was still alive by the end of 1969. Thus, the treatment of censoring may be important. We have also estimated the accelerated hazard model with the four other distributions discussed in Appendix C. It should be noted that the empirical distribution is known to be rising; hence, the exponential, which has a constant hazard, is a misspecification. The results using the other baseline hazards are quite similar. The four other distributions permit nonconstant hazards. These four functional forms have similar log likelihood values and likelihood ratios, though the loglogistic has a marginally better fit than the gamma with respect to the empirical hazard.

The normalized number of years of tobacco use has highly significant negative effects whose magnitudes are about 7 and 2 percent reduction in

TABLE 6.1

DORN SAMPLE: ACCELERATED TIME MODELS, WHOLE SAMPLE AND TWO COHORTS

Variable	All cohorts (t-statistic)	Cohort born before 1890 (t-statistic)	Cohort born 1891–1900 (t-statistic)
Intercept	6.92	6.88	6.83
	(2572.)	(1313.)	(1753.)
Proportion of years used tobacco occasionally	−.0214	−.0260	−.00975
	(−4.36)	(−3.29)	(−1.88)
Proportion of years used tobacco regularly	−.0688	−.0393	−.0515
	(−34.8)	(−11.4)	(−24.2)
Region			
South	−.00886	−.00630	−.00661
	(−4.54)	(−1.95)	(−2.98)
Northeast	−.00390	−.00386	−.00455
	(−2.07)	(−.123)	(−2.14)
North central	−.00434	−.00545	−.00423
	(−2.27)	(−1.70)	(−1.98)
Activity index	.00237	.000133	.00220
	(2.42)	(.0792)	(1.97)
Risk index	−.00288	−.00163	−.0000279
	(−5.65)	(−1.83)	(−.047)
Scale	.125	.107	.164
	(115.)	(107.)	(178.)
Shape	0.506	−0.774	−1.06
	(31.1)	(−13.6)	(−26.7)
LogL	−210299.	−42175.	−152951.

life span for regular smoking and occasional smoking, respectively. Thus, changing from not smoking to regular smoking would shorten a person's life noticeably. Information about the region where a person lived as of the mid-1950s generally indicates that life expectancies in the South, the Northeast, and the north central region were worse than in the West, but the magnitudes of the differences were small.

The physical activity index is also significant. Moving from an inactive 1 to an active 4 increases the time to death by about 0.7 to 1 percent. The occupational risk variable is highly significant though of small magnitude. The difference between a 1 and a 7, the lowest and highest risk groups, is about 1.4 to 2 percent shorter life.

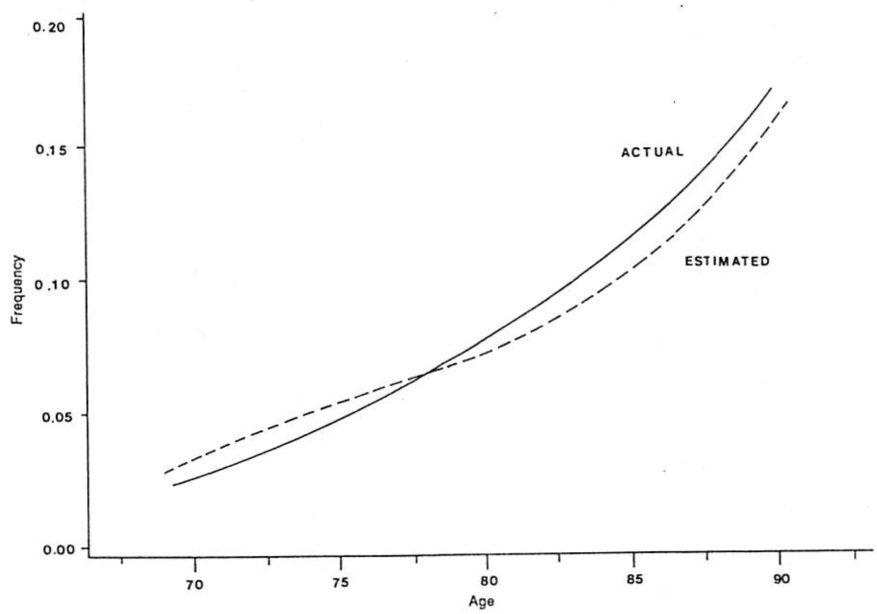
In Figure 6.1 we present the actual and estimated hazard based on the gamma function. The estimated hazard fits the data well with an R^2 of 0.892 and captures the increase in the hazard with age.

In the columns of Table 6.1 we present the accelerated hazard results with the Dorn sample divided into two age groups.[9] One is for 10,794 persons in the cohort born before 1891. (Normalized) regular tobacco use is highly significant and negative with a point estimate of about 4 percent per year used, whereas occasional use has a significant negative coefficient of about −2.5 percent. The southern inhabitants in the mid-1950s had significantly shorter life spans than western residents, but the differential is small.

Both the risk index and the physical activity index are based on the person's occupation in the mid-1950s when these men were in their mid-fifties or older. In these indexes a larger number means more activity and risk. The activity index is not significant. The risk index is almost significant in the gamma column; but the difference between 1 and 7 on this index is

FIGURE 6.1

DORN SAMPLE: ALL AGES, HAZARD FUNCTIONS, TIME-TO-FAILURE GAMMA FUNCTION



about 1 percent, so the effect is small. As risk increases, time to death is shortened.

The last column contains the results for the 59,151 sample members born between 1891 and 1900. The gamma distribution fits the data best and is the focus of our discussion. Each additional year of regular smoking reduces time to failure by a significant 5 percent (with the other specifications indicating slightly bigger impacts). The occasional smoking variable has a coefficient of about −1 percent, which is not significant.

Those from the South, the Northeast, and the north central region live less long than those from the West, though the differences are still not large. Increased physical activity associated with certain occupations leads to a significantly longer life (only for the gamma case), but the risk index is insignificant. All in all, the coefficients in the two right-hand columns of Table 6.1 are similar to each other and to those in the left-hand column, even though there are more truncated observations in the last column.

Table 6.2 contains the estimates using Cox's partial likelihood model,

TABLE 6.2

DORN SAMPLE: COX PARTIAL LIKELIHOOD ESTIMATES FOR HAZARD FUNCTION FOR DIFFERENT COHORTS

Variable	*Cohort born before 1890, estimate (t-statistic)*	*Cohort born 1890–1899, estimate (t-statistic)*	*All cohorts, estimate (t-statistic)*
Proportion of years used tobacco occasionally	0.287 (2.91)	0.157 (2.30)	0.212 (4.44)
Proportion of years used tobacco regularly	0.550 (12.6)	0.632 (28.2)	0.655 (33.4)
Region			
South	0.112 (2.91)	0.0853 (3.84)	0.0872 (4.69)
Northeast	0.0759 (2.01)	0.0674 (3.16)	0.0461 (2.57)
North central	0.0717 (1.84)	0.0626 (2.91)	0.0474 (2.61)
Activity index	0.0135 (0.67)	−.0141 (−1.28)	−.0173 (−1.85)
Risk index	0.0108 (1.01)	0.0107 (1.87)	0.0234 (4.87)
LogL	−55972.	−226841.	−311672.

which has an unknown baseline hazard and which estimates the effects of covariates on shifts in this baseline. The parameters shown indicate the percentage shift in the hazard rate with respect to a change in *X*. In this model an increase in the hazard rate is considered bad. Here we see that cigarette smoking occasionally and regularly are both associated with a higher hazard, with regular use about three times as bad.

The residents of the West have better hazards than denizens of other areas, with the worst area being the South; this set of results is roughly consistent with those in Table 6.1. The risk index is highly significant, with those in riskier occupations having a higher hazard. The physical activity index indicates that the more active have lower hazards though the coefficient is not significant at conventional levels.

In Figure 6.2 we present the actual and estimated proportional hazard. The baseline hazard is the one for the people who are in the West and do not smoke (the omitted categories). Normally the baseline category cannot be calculated, but we have a large enough sample to do so. Again the model fits the data well with an R^2 of 0.93.

The two right-hand columns of Table 6.2 contain the Cox estimates for the aforementioned cohorts. These estimates can be described quickly, since

FIGURE 6.2

DORN SAMPLE: ALL AGES, PROPORTIONAL HAZARD

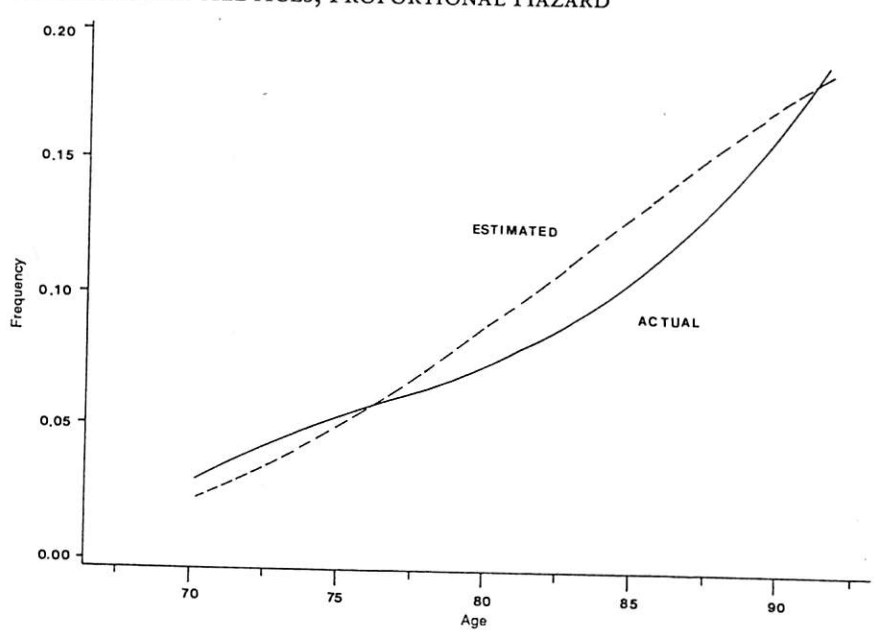


Table 6.3
RHS Sample:
Accelerated Time-to-Failure Models

Variable	Exponential, Estimate (t-statistic)		Weibull, Estimate (t-statistic)		Lognormal, Estimate (t-statistic)		Loglogostic, Estimate (t-statistic)		Gamma, Estimate (t-statistic)	
Intercept	7.48	7.42	6.75	6.74	6.72	6.72	6.72	6.72	6.71	6.70
	(84.7)	(84.1)	(1525)	(1456)	(1544.7)	(1522)	(1497.)	(1463.)	(1283.)	(1344.)
Black	−.136	−.115	−.0112	−.00930	−.0124	−.0114	−.0133	−.0124	−.0118	−.0102
	(−1.61)	(−1.39)	(2.69)	(−2.14)	(−2.98)	(−2.69)	(−3.09)	(−2.80)	(−2.91)	(−2.58)
Married	1.99	1.601	.124	.0914	.105	.0837	.112	.0855	.0984	.0829
	(24.2)	(22.2)	(26.8)	(22.1)	(27.9)	(26.6)	(26.9)	(24.4)	(26.5)	(29.9)
Widowed	−.477	−.668	−.0354	−.0509	−.0369	−.0478	−.0371	−.0517	−.0353	−.0384
	(−8.14)	(−12.2)	(−12.1)	(−17.1)	(−11.8)	(−15.6)	(−12.0)	(−17.1)	(−10.9)	(−11.2)
Divorced/ separated	1.233	1.24	.0742	.0781	.0755	.0765	.0745	.0769	.0766	.0779
	(5.65)	(5.70)	(6.78)	(6.79)	(8.93)	(8.95)	(7.96)	(7.97)	(9.94)	(10.9)
Education	.0115	.0110	.000331	.00394	.000352	.000367	.000381	.000435	.00031	.000213
	(1.47)	(1.45)	(.895)	(.985)	(.951)	(.973)	(1.00)	(1.12)	(.858)	(.609)
Social Security benefits (thousands)	.113	.113	.0124	.0127	.0123	.0122	.0132	.0132	.0118	.012
	(6.66)	(6.66)	(14.1)	(13.8)	(16.0)	(15.6)	(15.8)	(15.2)	(16.4)	(16.7)
Spouse working	−.579	—	−.0512	—	−.0345	—	−.0410	—	−.0261	—
	(−10.2)	(—)	(−17.5)	(—)	(−11.6)	(—)	(−13.3)	(—)	(−7.70)	(—)
Supplemental Security income (thousands)	.860	.977	.0652	(.0811)	.0559	.0605	.0624	.0715	.0518	.0488
	(3.34)	(3.74)	(4.63)	(5.32)	(5.70)	(5.98)	(5.01)	(5.40)	(5.95)	(6.03)

Number of dependents	−.0460	−.0686	−.00589	−.00747	−.00557	−.00660	−.00527	−.00662	−.00536	−.00572
	(−1.30)	(−1.95)	(−3.56)	(−4.32)	(−2.87)	(−3.36)	(−2.76)	(−3.47)	(−2.68)	(−2.82)
Pension income (thousands)	.0829	.0841	.00614	.00720	.00615	.00539	.00639	.00639	.00572	.00491
	(4.32)	(4.31)	(6.46)	(6.80)	(7.16)	(6.72)	(6.86)	(6.51)	(7.33)	(7.01)
Longest occupation										
Professional	.0935	.0927	.00683	.00536	.00683	.00679	.00731	.00746	.00639	.00611
	(1.31)	(1.30)	(1.91)	(1.42)	(2.01)	(1.97)	(1.98)	(2.05)	(1.97)	(1.96)
Management	.0536	.0669	.00431	.00515	.00781	.00819	.00706	.00753	.00853	.00894
	(.751)	.934	(1.12)	(1.37)	(2.25)	(2.32)	(1.98)	(2.06)	(2.54)	(2.75)
Scale	1.0	1.0	.0499	.0524	.0697	.0712	.0394	.0407	.0761	.0802
	(—)	(—)	(49.7)	(49.3)	(53.1)	(52.9)	(49.3)	(49.0)	(40.3)	(51.4)
Shape	—	—	—	—	—	—	—	—	.466	−.912
									(4.07)	(−8.74)
LogL	−3516.6	−3567.87	387.0	239.9	501.24	431.56	455.24	364.64	510.59	475.41

the parameters are similar to those obtained in the full sample in each cohort except that the activity and risk indexes are no longer significant (and in one case change signs). The results in this table are in qualitative accord with those in Table 6.1.

We now turn to the RHS sample. Table 6.3 contains the results for men in the RHS for the period 1969–1977 using accelerated hazard models. There are two covariate specifications for each distribution, with spouse working excluded in the second version. Note that in this sample, the censoring rate is nearly 80 percent. Nevertheless, all but the exponential distribution fit the data about equally well, but the gamma fits best. It appears that in this random sample blacks (about 9 percent of the sample) die about 1 percent younger. In comparison with single men, married and divorced men live somewhat longer, with coefficients about seven to twelve times as large as the black coefficient. Widowers have a shorter life span. Both the divorced and widower results are surprising given previous results on health (see, for example, Rosen and Taubman 1984).

Perhaps most interesting is that income greatly matters, with Social Security and pension benefits having about the same effect of 1 to 0.5 percent, though the former is somewhat larger. Social Security benefits are a function of an average of monthly earnings generally calculated over 20–25 years. The Social Security's transformation is highly regressive. Pension benefits are a function of plan used, with some plans based on lifetime contributions and pension earnings and others based on earnings in the last one to five years. Pensions are either proportional to the earnings base used or are progressive when a company integrates its pension plan with Social Security benefits. Although it is possible that chronic health problems affect earnings throughout a person's lifetime, it is difficult to believe that our findings on the two-income sources represent the common effect of poor health on income, given the huge differences in the way earnings are translated into Social Security and pension benefits. Supplemental Security Income also matters to a similar extent. This in some ways is surprising, since the variable is zero until we first measure it in 1975 and since some people who would have been eligible if they had lived until 1975 are classified as not eligible because they died earlier.

A working spouse has a negative effect of 3 percent in the gamma column, and this effect ranges up to 5 percent. This variable is statistically significant for all four distributional assumptions. It is possible that a wife at work cannot urge a husband to see a doctor, keep him out of bars and away from unhealthly consumption, or provide regular meals. However, it is also possible that having a chronically ill husband induces a wife to work. Thus, we also present equations omitting this variable; changes in other coefficients are minor.

The coefficient on dependents at home (who are present in only 7 percent of the cases) is a highly significant −0.6 percent. Perhaps such dependents require time and emotional resources that could otherwise be used to lengthen the caretaker's life. Alternatively, dependents may more likely be at home in cases in which morbidity experience suggests a higher hazard rate.

The RHS also records the longest occupation. We use this information directly in dummy variable form. (We omit the physical activity transformation that is highly colinear with the dummies used.) Compared to everyone else, managers and professionals live 0.6 to 0.9 percent longer.

In Figure 6.3 we present the actual and estimated hazards using the gamma distribution. The model fits the data very closely with an R^2 of 0.981.

Table 6.4 contains the hazard function estimates for the RHS sample based on Cox's (1975) partial likelihood model. In this model the married and widowed variables have opposite signs and are highly significant. Pen-

FIGURE 6.3

RHS HAZARD FUNCTIONS: TIME-TO-FAILURE GAMMA FUNCTION

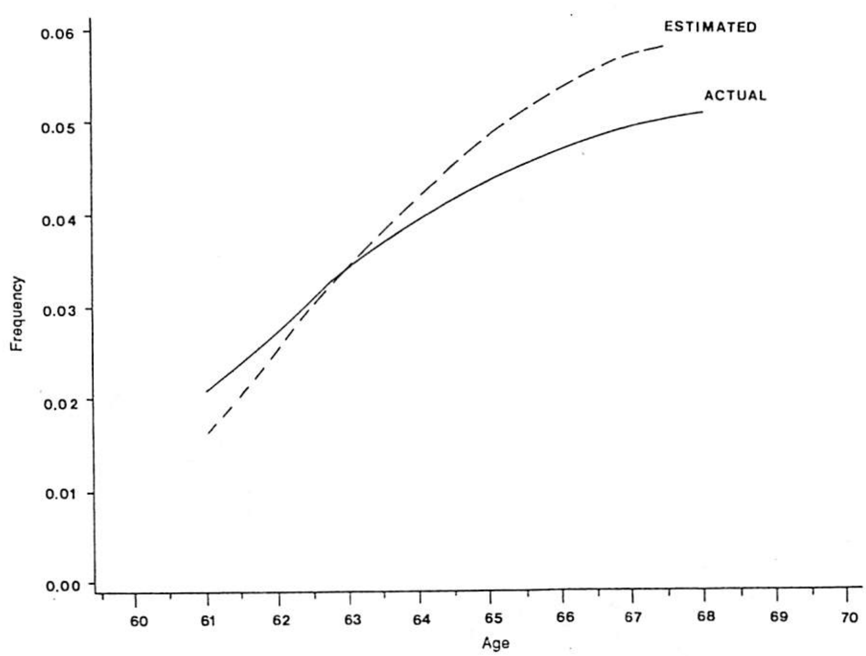


TABLE 6.4

RHS SAMPLE: PROPORTIONAL HAZARD MODEL (COX'S PARTIAL LIKELIHOOD)

VARIABLE	PROPORTIONAL HAZARD, ESTIMATE (T-STATISTIC)	
Intercept	—	—
	(—)	(—)
Black	.211	.165
	(2.53)	(1.99)
Married	−2.42	−1.74
	(−29.2)	(−24.3)
Widowed	.695	.952
	(11.9)	(17.3)
Divorced/separated	−1.46	−1.46
	(−6.71)	(−6.70)
Education	−.0066	−.0075
	(−.884)	(−1.01)
Social Security benefits	−.238	−.229
(thousands)	(−13.4)	(−13.0)
Spouse working	.960	—
	(16.6)	(—)
Supplemental Security income	−1.25	−1.47
(thousands)[a]	(−4.50)	(−5.14)
Number of dependents	.110	.137
	(3.30)	(4.15)
Pension income	−117.	−.128
	(−6.15)	(−6.4)
Longest occupation		
Professional	−.121	−.092
	(−1.69)	(−1.2)
Management	−.0909	−.09
	(−1.27)	(1.35)
LogL	−13533.2	−13663.6

a. Measured in 1975.

sion and Social Security benefits are of roughly the same size and are highly significant and negative, as are the supplemental security benefits. Dependents at home raise the hazard rate. Figure 6.4 displays the observed and estimated hazards. The fit is good with an $R^2 = 0.822$.

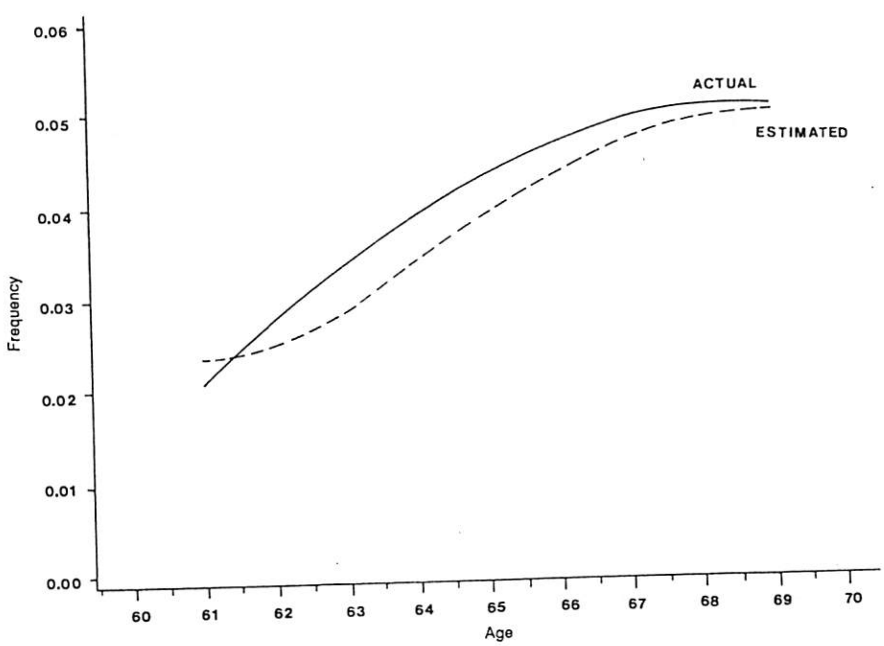
Conclusion

In this paper we have used the RHS and the Dorn samples to examine the relationship between mortality and various sociodemographic and life-style measures. We do so using time-to-failure and hazard models. These models are made more complicated because 60–80 percent of the people were still alive at the end of the sample frames.

To surmount censoring we have used a variety of assumptions in our time-to-failure models. We find that these models can be fit to the data even with the large amount of censoring and a relatively large number of covariates, some of which vary over the sample period.

In the time-to-failure models, the exponential function implies a constant hazard rate that is inappropriate for mortality data. The other four distributions usually yield similar results among them, but in a few cases the impact of a variable differs depending on the distributional assumption used. In terms of goodness of fit, the gamma and loglogistic seem best for the baseline hazard.

Figure 6.4
RHS Proportional Hazard



We find that regular tobacco users have shorter lives and that there are regional variations and occupational differences. In the RHS, married and divorced men live longer than those widowed or never married. We also find that pensions and Social Security benefits, both of which are related to preretirement work, have strong effects on life expectancy with Social Security having somewhat stronger effects. We also find that having a child at home or a working spouse when a man is in his sixties is associated with shorter life spans, though in both cases the direction of causality is questionable. Those in the professional and managerial occupations live longer. Education, however, is not statistically significant.

These are preliminary results based on fewer observations and more limited assumptions than we will eventually be able to utilize. It will be of interest to see how robust these findings are in our future analyses.

APPENDIX A

DORN SAMPLE: SUMMARY STATISTICS FOR DIFFERENT COHORTS

Variable	COHORT BORN BEFORE 1890 Mean	S.D.	COHORT BORN 1890–1899 Mean	S.D.	ALL COHORTS Mean	S.D.
Proportion of years used tobacco occasionally[a]	.0330	.142	.0318	.143	.0307	.138
Proportion of years used tobacco regularly[a]	.495	.327	.546	.353	.543	.343
Region						
South	.258	.438	.240	.425	.249	.430
Northeast	.300	.458	.307	.461	.290	.452
North central	.264	.441	.289	.452	.271	.442
Activity index	2.17	.848	2.13	.830	2.19	.841
Risk index	2.01	.158	.199	.156	2.20	1.70
Number of observations	10784		59147		85627	
Percentage censored	40.04		63.7		65.5	

a. Number of years of tobacco use divided by age at survey date minus ten.

APPENDIX B

RHS SAMPLE: SUMMARY STATISTICS

Variable	Mean	S.D.
Black	0.874	.284
Married	0.585	.493
Widowed	0.131	.338
Education	9.84	3.52
Social Security benefits (thousands)	1.29	1.28
Spouse's earnings (thousands)	1.47	3.08
Supplementary Security income (thousands)	.0245	.168
Dependents	.143	.560
Pension income (thousands)	1.31	2.52
Professional	.179	.383
Management	.161	.367
Number of observations	7915	
Percentage censored	78.3	

In the proportional hazard model, covariates act multiplicatively on the hazard. If t is the random variable denoting failure time, then

$$\lambda(t;x) = \lambda_o(t)\, e^{\{x\beta\}}$$

where $\lambda_o(t)$ is the baseline hazard with an unspecified distribution. The survivor function is given by:

$$S(t;x) = [exp\{-\int_o^t \lambda_o(\xi)d\xi\}]^{e^{\{x\beta\}}}$$

where exp is the exponential function.

The accelerated hazard model specifies $\log T$ as a linear function of the covariates, that is, $y = \log T = x\beta + \sigma\varepsilon$ where ε is a random disturbance. Alternatively, failure time can be expressed as $T = exp\{x\beta\}T_o^{\sigma}$ where $T_o = \exp\{\sigma\varepsilon\} > 0$ has a baseline hazard function $\lambda_o(t)$ that is independent of β. The baseline hazard function depends on $te^{\{-x\beta\}}$. In terms of the baseline hazard, the hazard for T is:

$$\lambda(t;x) = \lambda_o e\{-x\beta\}$$

with survivor function

$$S(t;x) = exp[\, - \int_o^t e^{\{-x\beta\}}\, \lambda_o(\xi)d\xi].$$

The covariates act multiplicatively on failure time rather than on the hazard. That is, they change the rate at which an observation moves along the baseline hazard over time and thus accelerate or decelerate the time to failure.

Our estimates of the proportional hazard model are based on the maximization of the partial likelihood function (Cox 1975). The nuisance baseline hazard function is not specified. However, the resulting estimates of the covariate effects are still consistent and asymptotically normal. Because of the relatively large data bases we are examining, we are able to use observations from the omitted categories in the covariate list to estimate the baseline hazard with the Kaplan and Meier (1958) product-limit estimator. We use the same estimator when examining the degree to which the alternative models of failure time fit the underlying observed sample hazards.

Maximum likelihood estimates of the accelerated-failure-time model are based on five different specifications for the distribution of the baseline failure time: exponential, Weibull, lognormal, loglogistic, and gamma. The hazard functions for these different baseline distributions in the log-linear regression models are:

(exponential) $\lambda(t;x) = e^{\{-x\beta\}}$;

(Weibull) $\lambda(t;x) = \sigma^{-1} e^{\{-x\beta/\sigma\}} t^{(\sigma^{-1}-1)}$;

(lognormal)
$$\lambda(t;x) = [\, exp\{ -(log(t)-x\beta)^2/2\sigma^2\}/(\sigma\, t\sqrt{2n})]/$$
$$[\textstyle\int^{\infty} f(\xi)d\xi\,]$$
$$(ln(t)-x\beta)/\sigma$$

where $f(\xi)$ is the standard normal density function;

(loglogistic) $\lambda(t;x) = e^{\{-x\beta/\sigma\}}\sigma^{-1}t^{(\sigma^{-1}-1)}/(1 + e^{\{-x\beta/\sigma\}}t^{-\sigma})$;

(gamma)
$$\lambda(t;x) = [|\gamma|(t^{\gamma}/\gamma^2)^{(1/\lambda^2)}e^{\{-t^{\gamma}/\gamma^2\}}\,(t\Gamma(1/\gamma^2)]/$$
$$[\delta + (-1)^{\delta}\Gamma(1/\gamma^2, t^{\gamma}/\gamma^2/\Gamma(1/\gamma^2)]$$

where $\delta = [{}^{1 \text{ if } \gamma<0}_{0 \text{ if } \gamma>0}\,]$ and where $\Gamma(\alpha_1)$ is the incomplete gamma function, $\Gamma(\alpha_1, \alpha_2)$ is the incomplete gamma function, and γ is the gamma distribution shape parameter.

Discussion

Richard V. Burkhauser

This interesting paper provides the first results from a relatively new data set, new at least for economists—the Dorn Sample of Veterans. The paper uses hazard modeling to estimate a baseline survivor distribution for both the Dorn sample and for men in the Retirement History Survey. This technique allows us to measure the risk of death over time for an age cohort while estimating covariates that control for cohort heterogeneity.

The paper is of interest both as an attempt to use a relatively new statistical technique—hazard modeling—and because it introduces economists to a new and potentially important data set. Several specific issues can be raised with respect to the limits of the data and the inferences of the model. However, they should be prefaced by the old saying that: "Upon seeing a talking horse with a Southern accent, the wonder is not how he got the accent but that he talks at all." Particularly in discussing the limits of the Dorn data, it must be recognized that the authors have taken on an enormous task in analyzing this data set, and I look forward to how additional economic inferences will be drawn from its limited variables.

As the authors recognize, researchers in demography, statistics, and engineering among other disciplines use hazard modeling to a much greater extent that economists. They are usually interested in the timing of events like death, the breakdown of a machine, or the movement of a star. For the most part, it is only to control for heterogeneity that covariates are introduced. Hence, they spend little time or effort on developing a behavioral model linked to such covariates or in making inferences about their values. To this point, most economists who have used hazard models have also emphasized the time parameters and the problem of separating true time dependence from observed or unobserved heterogeneity. Heckman and Singer (1984b), for instance, offer one solution to this problem. Flinn and Heckman (1983) in their early work model the hazard of return to work and see if people who say they are out of the labor force do, in fact, stay out of work longer than those who say they are unemployed. Little is made of the covariates used in the model.

A simple example of the separation problem can be found in Holden, Burkhauser, and Feaster (1987). Workers are observed at the start of retirement and followed thereafter so that their risk of falling into poverty can be estimated over time. A hazard model is used to see if this risk increases or decreases. If we simply follow the experience of this cohort, a strong negative time dependence is found. That is, the risk of falling into poverty falls over time. But this may be caused by heterogeneity within the cohort. The amount of wealth held at retirement will also affect the speed of falls into poverty. The negative time dependence may simply be an artifact. It may simply be the case that in the first period those closest to the poverty line are more likely to become poor and hence leave the sample. In the second period, the survivors who were initially wealthier and hence less likely to fall into poverty remain. This wealthier cohort will have a lower risk of falling into poverty in the next period. Unless this cohort heterogeneity is controlled, it will be confused with negative time dependence.

Depending on the question asked, we may or may not need to control for heterogeneity within the cohort. A more subtle issue concerns controlling for unobserved heterogeneity, or the so-called nuisance parameter, and separating its effects from true time dependence. Unobserved variation within the cohort may effect the timing of the risk. For instance, in the unemployment literature is it "scaring"—that is, time in unemployment—that decreases the likelihood of finding employment, or is it "unobserved heterogeneity"—that is, that people with a strong but unobserved work ethic quickly get jobs, leaving the "loafers" behind? In this paper, issues of unobserved heterogeneity are not considered. Rather, the model is fitted to various specific functional forms and the major discussion is on the value of the covariates. There is little discussion of the economic model these covariates are testing.

The authors point out a major advantage of a hazard model—it addresses the problem of incomplete spells or right censoring. That is, a hazard model allows the use of information on individuals in a sample who have not encountered the hazard—such as death, poverty, or retirement—by the time the sampling period ends. Hence their spell of life, nonpoverty, or work is incomplete. Because most people have not died in both the data sets tested, this is a compelling reason for using such models. There is little discussion, however, of left censoring. That is, both data sets first observe people in the middle of a spell (in the middle of life). The RHS first follows men at ages 58–63. In the Dorn study, the ages vary but anyone not alive in 1954 is excluded from the sample. Again, this may or may not be a problem, depending on the specific questions the researchers are addressing. For instance, if they are interested in how occupations affect mortality rates but pick up their sample at advanced ages, they select out those who have died previously. If the risk associated with the disease is quick and fatal (for example, working the night shift at a 7–11 store), this is a major problem. If it is slow and only shows up at much older ages (such as black lung), it is less of a problem. But it is still a problem to the degree that people who work in hazardous jobs have adjusted investment in health to compensate for the reduction in health associated with a specific occupation.

The major finding in the Dorn study is that positive time dependence exists. That is, the risk of dying increases over time. In addition, smoking increases the risk of death at any age. This seems perfectly reasonable. It is in measuring the magnitudes of the effect that a fuller discussion of the problem of left censoring must be undertaken. This is particularly true of the region variables and the occupation variables. How useful are measures of region or occupation at advanced ages in discussions of mortality? This must greatly depend on the mobility of people across regions and their occupations at younger ages. This is particularly a problem in the subsample of veterans who in 1953 were aged 60 and over. At these ages it is likely that many of the veterans are already retired from their career jobs.

The RHS data is far richer in socioeconomic variables than the Dorn data, although the data are for a much shorter time period, 1969–1979. The findings here are once again reasonable. The risk of dying increases over time. The authors recognize the endogeneity of labor supply decisions and have developed a model elsewhere; see Sickles and Taubman 1986, in which health and work decisions are considered simultaneously. Butler, Anderson, and Burkhauser (1987) test a bivariate hazard model which shows that health decisions and decisions about re-entry into the labor market after retirement are made simultaneously.

Here the authors attempt to avoid the issue by explicitly excluding labor supply variables. This strategy is not fully successful, however. The authors calculate the earnings record of workers only up to 1969. But they then estimate benefits available in each year thereafter. These yearly benefits are positively correlated with age and time. The actuarial adjustment of Social Security between ages 62 and 65 increases the yearly benefit by 25 percent and hence is highly sensitive to age. In addition, the benefits themselves were increased by congressional action over this time period. Hence, there is a built-in positive correlation between the size of benefits and survival that confounds the results. Depending on what the Social Security variable is supposed to be measuring—that is, past earnings or current wealth—a better measure might be the Average Indexed-Monthly Earnings (AIME), a good proxy of lifetime earnings used by the Social Security Administration to calculate benefits. It is less sensitive to the age at which benefits are taken.

The next surprising finding is that SSI receipt means a lower risk of death. Since recipients of SSI are clearly at the lower end of the wealth distribution, this is a rather odd finding. Again this may be simply an artifact of the data. Since a person has to be age 65 to get SSI and the program was not started until 1974, when workers in this sample were aged 63–68, it may simply be picking up the fact that a worker could only get an SSI value in this data if he lived until the 1975 sample. A final caveat is in regard to the authors' use of R^2 to choose between the various models presented. Since the equations are nonnested, the R^2 statistic is inappropriate. A more appropriate measure is the Kolmegorov-Smirnov test.

Despite these reservations, this paper provides an interesting first run through the data. More time and effort by economists should go into considerations of the importance of covariates as well as into the difference in time dependence across groups. The preliminary results are not enormously surprising, but we would have less faith in the data if they were. The risk of death does increase over time. Cigarette smoking is a sure way of avoiding the problems of old age. The authors are well aware of most of the issues raised here, and I suspect the next time out this horse will not only talk but maybe even tap dance.

Notes

1. See Sickles and Taubman (1986) and references therein.

2. See, for example, Rosen and Taubman (1984), Kitagawa and Hauser (1973), and Duleep (1986) and references therein.

3. See, for example, Cox (1972), Manton, Stallard, and Vaupel (1986), Kalbfleisch and Prentice (1980), and Vaupel, Manton, and Stallard (1979). These techniques have been used by economists on different problems such as the duration of unemployment. See, for example, Heckman and Singer (1984) and Lancaster (1979).

4. This information on date and cause of death has recently been assembled through 1983, but we have not yet had access to information for years after 1979.

5. For details, contact Z. Hrubec, Radiation Epidemiology Branch, National Cancer Institute, Landow Building, Room 3A22, Bethesda, Maryland 20892.

6. Separation by cause has a number of advantages. First, we can determine the influence of expenditures in various National Endowment for the Humanities research programs on specific diseases and determine their separate price responsiveness. Such information should help in allocating research budgets. Second, the data on death may yield biased results that arise because of aggregation, especially if functions are not linear. Third, better decisions on both research and treatment facilities can be made if we have advance knowledge on what diseases are likely to strike the elderly, who will be a much larger share of the population in the future. Fourth, the hazard functions by cause of death may more confidently be extrapolated to older ages to predict what will happen to currently elderly people, though it is possible that this is offset by negatively correlated misclassification error.

7. The sample has only a small number of nonwhites.

8. We thank Richard Burkhauser and Laurence Kotlikoff for these suggestions.

9. We do not include separate results for the youngest age group since only 5 percent died by 1969. This age group is included in all the cohort groups.

References

Anderson, K., and R. Burkhauser. 1983. "The Effect of Actual Mortality Experience Within a Retirement Decision Model." Mimeo. Nashville, Tenn.: Vanderbilt University.

Beebe, G., and A. Simon. 1969. "Ascertainment of Mortality in the U.S. Veteran Population." *American Journal of Epidemiology* 89:636–43.

Butler, T., K. Anderson, and R. Burkhauser. 1987. "Work and Health After Retirement: A Bivariate Interval Hazard Model Approach." Working Paper, Vanderbilt University.

Cohen, B. 1953. "Methodology of Record Follow-up Studies on Veterans." *American Journal of Public Health* 43:1292–98.

Cox, D. 1972. "Regression Models and Life Tables." *Journal of the Royal Statistical Society* (Series B) 33:187–202.

———. 1975. "Partial Likelihood." *Biometrika* 62:269–76.

DeBakey, M., and G. Beebe. 1952. "Medical Follow-up Studies on Veterans." *Journal of the American Medical Association* 182:1103–9.

Dorn, H. 1958. "The Mortality of Smokers and Non-Smokers." In *Proceedings, Social Statistics Section, American Statistical Associations,* pp. 34–71.

Duleep, H. 1986. "The Socioeconomic Determinants of Mortality: The Role of Income." *Journal of Human Resources* 21:238–51.

Flinn, C., and J. Heckman. 1983. "Are Unemployment and Out of the Labor Force Behaviorally Distinct Labor Force States?" *Journal of Labor Economics* 1:28–42.

Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80:223–55.

Heckman, J., and B. Singer. 1984a. "Econometric Duration Analysis." *Journal of Econometrics* 24:63–132.

———. 1984b. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52:271–320.

Holden, K., R. Burkhauser, and D. Feaster. 1987. "The Timing of Fall into Poverty After Retirement: An Event History Approach." Working Paper, Vanderbilt University.

Kahn, H. 1966. "The Dorn Study of Smoking and Mortality Among U.S. Veterans: Report of 8 1/2 Years of Observation." In W. Haenszel, ed., *Epidemiological Approaches to the Study of Cancer and Other Diseases*. Monograph 19. Bethesda, Md.: U.S. Public Health Services, National Cancer Institute.

Kalbfleisch, J., and R. Prentice. 1980. *Statistical Analysis of Failure Time Data*. New York: Wiley.

Kaplan, E. L., and P. Meier. 1958. "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association* 53:457–81.

Kitagawa, E., and P. Hauser. 1973. *Differential Mortality in the United States of America: A Study in Socioeconomic Epidemiology*. Cambridge, Mass.: Harvard University Press.

Lancaster, T. 1979. "Econometric Methods for the Duration of Unemployment." *Econometrica* 47:939–56.

Manton, K. G., E. Stallard, and J. W. Vaupel. 1985. "Alternative Models for the Heterogeneity of Mortality Risks Among the Aged." *Journal of the American Statistical Association* 81:635–44.

Rogot, E. 1974. *Smoking and General Mortality Among U.S. Veterans 1954–1969.*

DHEW Publication (NIH) 74-544. Washington, D.C.: Department of Health, Education and Welfare.

Rogot, E., and J. Murray. 1980. "Smoking and Causes of Death Among U.S. Veterans: 16 Years of Observation." *Public Health Reports* 95:213–22.

Rosen, S., and P. Taubman. 1984. "Changes in Impact of Education and Income on Mortality in the U.S." In *Statistical Uses of Administrative Records with Emphasis on Mortality and Disability Research*. Washington, D.C.: U.S. Department of Health, Education and Welfare.

Sickles, R., and P. Taubman. 1986. "An Analysis of the Health and Retirement Status of the Elderly." *Econometrica* 54:1339–56.

Taubman, P., and S. Rosen. 1982. "Healthiness, Education and Marital Status." In V. Fuchs, ed., *Economic Aspects of Health*. Chicago: University of Chicago Press.

Vaupel, J., K. Manton, and E. Stallard. 1979. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16:439–54.