Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study
Author(s): Michael D. Hurd and Kathleen McGarry
Source: *The Journal of Human Resources*, Vol. 30, Special Issue on the Health and Retirement Study: Data Quality and Early Results (1995), pp. S268-S292
Published by: University of Wisconsin Press
Stable URL: http://www.jstor.org/stable/146285
Accessed: 05-07-2016 20:34 UTC

Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study

Michael D. Hurd
Kathleen McGarry

ABSTRACT

In the Health and Retirement Study respondents were asked about the chances they would live to 75 or to 85. We analyze these responses to determine if they behave like probabilities of survival, if their averages are close to average probabilities in the population, and if they have correlations with other variables that are similar to correlations with actual mortality outcomes. We find that generally they do behave like probabilities and that they do aggregate to population probabilities. Most remarkable, however, is that they covary with other variables in the same way actual outcomes vary with the variables. For example, respondents with higher socioeconomic status give higher probabilities of survival, whereas respondents who smoke give lower probabilities. We conclude that these measures of subjective probabilities have great potential use in models of intertemporal decision-making under uncertainty.

I. Introduction

Many economic models are based on forward-looking behavior by economic agents. Although it is often said that "expectations" about future events are important in these models, more precisely it is the probability distributions of future events that enter the models. For example, an individual's decisions about consumption and saving are thought to depend on beliefs about future

Michael D. Hurd is a professor of economics at the State University of New York and a researcher at the National Bureau of Economic Research. Kathleen McGarry is a professor of economics at the University of California–Los Angeles and a researcher at the National Bureau of Economic Research. *Financial support from the University of Michigan and the National Institute on Aging is gratefully acknowledged. Many thanks to Nancy Cole for research assistance. The data used in this article are from the alpha release of the HRS.*

interest rates, the likelihood of dying, and the risk of substantial future medical expenditures. According to this theory, decision makers have subjective probability distributions about these and other events and they use them to make decisions about saving.

In a few microeconomic models, we have data on probability distributions that may plausibly be assumed to approximate those required by the models of decision-making under uncertainty. For example, in life-cycle models of consumption, mortality risk influences saving. These kinds of models have been estimated by assuming that individuals have subjective probability distributions on mortality risk that are the same as those found from life tables (Hurd 1989).

In most applications, however, we do not have data on probability distributions, so estimation requires some unverifiable assumptions. For example, in macroeconomic models expectations are assumed to be rational, which often yields an estimable relationship; yet, the rationality assumption cannot be tested outside the context of the model. In life-cycle models of saving, the average subjective mortality risk of a cohort may not be well approximated by life-table mortality rates because a cohort may not believe that the mortality experience of older cohorts will be the same as theirs due to projected improvements in mortality rates. Furthermore, individuals within a cohort will have different subjective probability distributions on mortality risk because of observable and unobservable differences in mortality risk factors. Finally, an individual's own subjective evaluation of probability distributions determines behavior, even if it is systematically incorrect; yet that evaluation is not generally observable.

The Health and Retirement Survey (HRS) contains a number of innovative questions about the chances of future events such as working full-time past age 62 or living to age 75.[1] Respondents were asked to indicate on a scale of 0–10 the chances of such events. After rescaling to 0–1, the responses can be interpreted as subjective probability distributions of the events. These kind of questions have the potential to change substantially the way in which we estimate stochastic dynamic models based on microdata because they can supply probabilities of events for which we have no population averages, and because they contain individual heterogeneity about probabilities. They can, in principle, be used directly in our models of decision-making.

Questions about expectations have been asked in other surveys. For example, the Retirement History Survey (RHS) asked respondents about the age at which they expected to retire, about what they expected to receive in Social Security benefits, and about their expectations of postretirement income. Some of these responses have been analyzed by Hall and Johnson (1980), Anderson, Burkhauser, and Quinn (1986), and most extensively by Bernheim (1988, 1989, and 1990). Bernheim identified two important limitations of questions about expectations. First, "expectation" may not have a well-defined meaning in a survey: with respect to the age of retirement respondents seemed to think of the mode

1. See Juster and Suzman (1993) for a description of the HRS.

or most likely age of retirement rather than the average outcome, which would be the empirical manifestation of the mathematical mean or expectation; however, with respect to Social Security benefits, the response was close to the average outcome. Thus, in general there could be a difference between the analyst's interpretation of a question about expectations and the respondent's interpretation. Second, the rate of item nonresponse was large: for example, in Bernheim's study of the evolution of expected Social Security benefits, his data set was limited to 370 observations, even though the RHS questioned about 11,000 households at the initial interview (Bernheim 1990). There was some selection on personal characteristics, but the main cause was nonresponse to the question on expected Social Security benefits.

Using data from a small survey, Hamermesh and Hamermesh (1983) and Hamermesh (1985) studied the subjective probabilities of survival to 60 or to 80. They found the probabilities were reasonably consistent with life-table probabilities, and that the qualitative variation with risk factors was similar to what is found in epidemiological data. A quantitative comparison cannot be exact because their sample is not representative of the population. Nonetheless, their findings offer good evidence that the responses were reasonable, and they should encourage further investigation.

Dominitz (1993) analyzed data from wave 1 of the Survey of Economic Expectations about the subjective probabilities of unemployment, work, earnings, and income levels. The response rates to the questions about unemployment and work were high, about 96 percent. They were lower on the other questions, between 65 percent and 75 percent. The overall assessment of the probability questions is that they vary in a reasonable way with qualitative expectations. In further analysis by Dominitz and Manski (1994a) the median of the subjective distribution of future income was found to vary almost one-for-one with actual income. These results should increase our confidence that respondents on average understand probability questions and at least qualitatively answer them appropriately.[2]

While the HRS questions about subjective probabilities have great potential, it is certainly possible that, as an empirical matter, they are not particularly useful. For example, respondents may have little idea of the probabilities of future events, or they may answer at random. Of course, in panel data we can find if the probabilities correspond to actual outcomes, and whether they are related to behavior. However, even in cross-section data we can find if they have characteristics that make it plausible they will be useful.

The broad goal of this paper is to evaluate the subjective probability distributions of survival to 75 or 85. Our methods will be to compare the average probabilities with survival probabilities calculated from life tables; to study the internal consistency of the subjective probability distributions to see if they behave like probabilities; and to find if the probabilities vary qualitatively with observable risk factors as they do in actual outcomes.

2. See also Dominitz and Manski (1994b) for an analysis of subjective probabilities among high school students.

II. Measures of Subjective Probability Distributions in the HRS

The HRS contains a number of questions that can be interpreted as subjective probability distributions. We will study responses to the questions:

> "Using any number from zero to ten where 0 equals *absolutely no chance* and 10 equals *absolutely certain,* what do you think are the chances you will live to be 75 or more?"
>
> "85 or more?"

The questionnaire also included questions about the probability of working, of housing purchase, job stability, financial help to family, housing prices, Social Security, and the economy.

Respondents in face-to-face interviews were presented with a physical scale like the following:

00	01	02	03	04	05	06	07	08	09	10

Absolutely no chance — Absolutely certain

After normalizing to [0, 1] we treat the responses to the questions about living to 75 or 85 as measures of the subjective probabilities of survival, and call them P75 and P85. We have chosen to focus on these probabilities because much more is known about what constitutes reasonable answers than about the other subjective probabilities both with respect to their levels and to how they covary with other observable data.[3]

III. Probabilities of Living to 75 or 85

The HRS is a representative sample of individuals born in the years 1931–1941 (approximately 51–61 at interview), except that blacks, Hispanics, and Floridians were oversampled. For population comparisons, therefore, our sample is restricted to the age range 51–61, and we use sampling weights to account for the oversampling of blacks, Hispanics, and Floridians. For analysis we use a sample of men aged 51–65 and women aged 46–61 who were not represented by a proxy interview. We realize that outside of the age range 51–61, the sample is not representative of the population because an age-ineligible respondent must be a spouse of an age-eligible person. Nonetheless, we wanted more age variation than in the age-eligible sample, particularly because we want to find how the subjective probabilities vary as age approaches 75. Furthermore, about

3. See Hurd and McGarry (1993a and 1993b) for evaluation of the subjective probability of working past 62 or 65.

23 percent of the sample is outside the age range 51–61, which is a large fraction to drop in the absence of a compelling reason.

We have 7,946 observations that we will use in our analysis of P75.[4] (We have slightly fewer responses to P85.) The response rate in the entire survey to P75 and P85 is about 98 percent. This very high response rate by itself makes the use of subjective probabilities appealing.

A. *Comparisons with Life Tables*

We interpret an individual's subjective probability of surviving to 75 to be the probability of a success on a binomial random variable. In a population of n of a given age, the expected number of survivors to 75 would, therefore, be Σ(P75), and Σ(P75)/n would be the expected average survival rate. If the subjective probabilities are accurate on average, and if mortality risk is stationary, the average survival rate would be well approximated by the conditional survival probability calculated from a life table. Similarly, the expected number of survivors to 85 would be Σ(P85), and the expected number of survivors to 85 out of the survivors to 75 would be Σ(P85)/Σ(P75). That is, averages of the subjective probabilities act like survival probabilities and conditional survival probabilities calculated from a life table.

Table 1 contains the averages of P75 and P85 over ages 51–61 and, for comparison, a weighted average of the age-specific survival rates from life tables. The weights are the relative number of observations at each age by sex in the HRS.

On average the HRS respondents gave lower survival probabilities to 85 than to 75, so the implied conditional survival rate is less than 1.0, about 0.66. The levels of P75 averaged over men and women are close to the averages in the 1990 life table, but the P85 are higher than those from the life table. Taking the life table as the relevant comparison, men substantially overestimate the probability they will live to 85, and women underestimate the probability they will live to 75. As a consequence, both overestimate the conditional survival rate to 85 given survival to 75.

Mortality risk has declined over a number of years, and it is forecast to fall further. It seems reasonable that the HRS respondents would know of the trend in mortality and extrapolate continued improvements.[5] The second part of the table has information about the trends: it has probabilities of survival to 75 or 85 from age 55 calculated from a 1980 life table (based on observed age-specific mortality rates in 1980), from a 1988 life table, from a 1990 life table, and from a year-2000 life table (based on forecasts of changes in mortality risk).

The changes in survival probabilities are substantial, which makes it difficult to know what is a good standard of comparison to the HRS data: the 1990 life

4. The data come from the alpha release of the HRS, which has observations on 9,495 individuals. We dropped 474 observations mainly because the response was by proxy, in which case questions about subjective probabilities were not asked. We dropped an additional 913 observations because of age. In 162 cases the response to P75 was missing, and in 180 cases the response to P85 was missing.

5. Or they may simply feel more optimistic and fit than would a population that had not experienced any improvements in mortality.

Table 1
Average Probabilities of Living to Age 75 or 85

	Men		Women		All	
	To 75	To 85	To 75	To 85	To 75	To 85
HRS data[a]	0.62	0.39	0.66	0.46	0.65	0.43
1990 life table[a]	0.60	0.26	0.75	0.45	0.68	0.36
From age 55						
HRS data	0.64	0.40	0.67	0.46	0.66	0.43
1980 life table	0.54	0.21	0.73	0.41	0.64	0.31
1988 life table	0.59	0.24	0.74	0.42	0.67	0.33
1990 life table	0.60	0.25	0.75	0.44	0.67	0.35
2000 life table	0.62	0.28	0.78	0.51	0.70	0.40

Source: Authors' calculations from HRS and various life tables for the U.S.
Note: Based on 6,802 observations.
a. Ages 51–61 only.

table is the product of age-specific mortality rates in 1990, which could be quite different from the age-specific mortality risks the HRS population anticipates. For example, the averages of P75 and P85 over all 55-year-olds from the HRS are quite close to the probabilities from the year-2000 life table, whereas P85 is considerably larger than the probability from the 1990 life table. From this point of view, even the "overestimate" of P85 by men could be a reasonable projection.

Figure 1 has the distributions of P75 and P85. They have considerable bunching at 0, 0.5, and 1.0. An interpretation is that people choose one of the three points according to whether they are rather confident, not confident at all, or uncertain about living to 75 or 85. However, there are mini-spikes at 0.2 and 0.8, and particularly for P85, considerable mass at other points. A partial explanation for the bunching, particularly at 1.0 and 0, is that the scale offered to respondents was rather coarse: for example, a respondent with a subjective probability of 0.95 might round to 1.0.

In Figure 2, we have, for the moment, extended our sample to include men aged 46–74, and, in Figure 3, women aged 38–65.[6] We did this to get the greatest possible age range. As a reminder of the thin sample at ages far from the HRS age range, we show the distribution of observations at the bottom of the graph. The averages by age of P75 and P85 are compared with estimates of P75 and P85 from the 1990 life table. As we saw earlier P85 is considerably greater than the life-table estimates. What is most notable is that the age-paths of P75 and P85 are rather flat until about age 64 when they rise rather sharply. Figure 3 shows age-paths of P75 and P85 for women. The paths are flat and possibly even declining before 60.

6. We only use the extended sample for these two graphs.

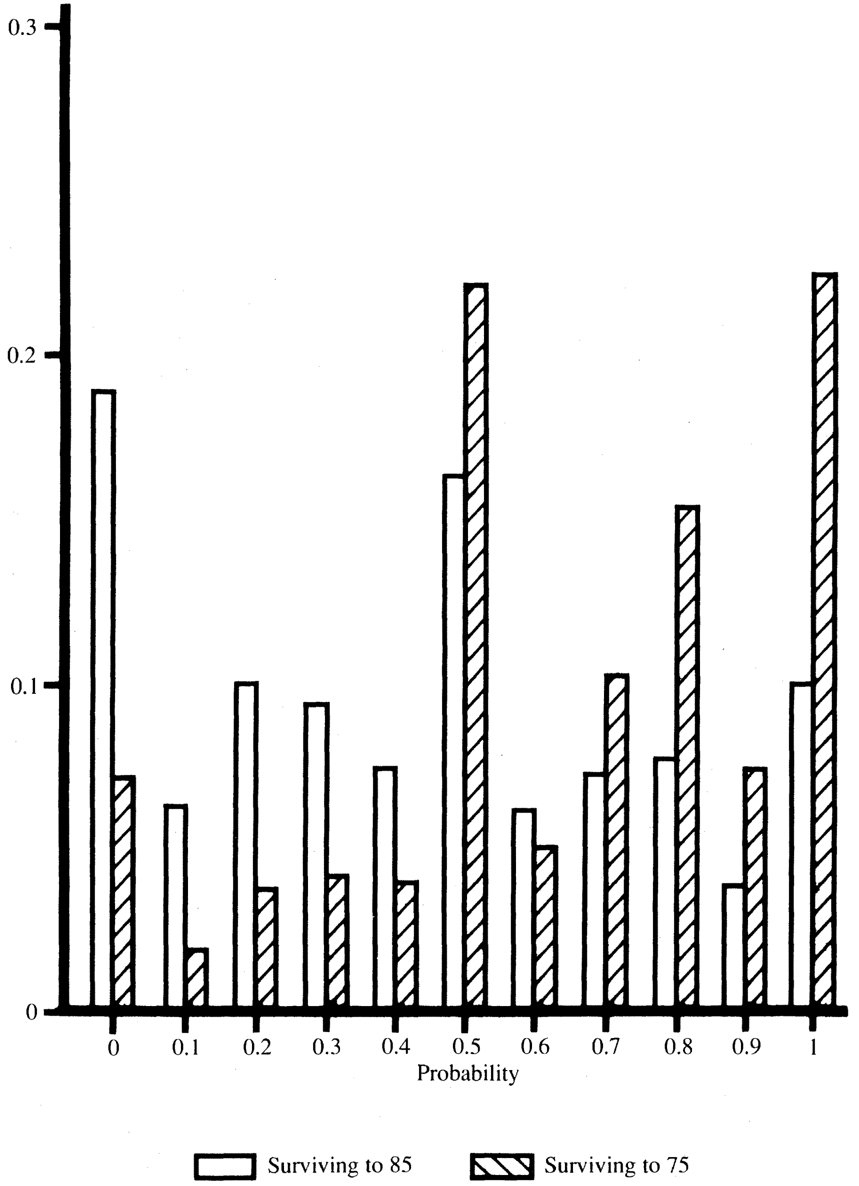

Figure 1
Distributions of survival probabilities

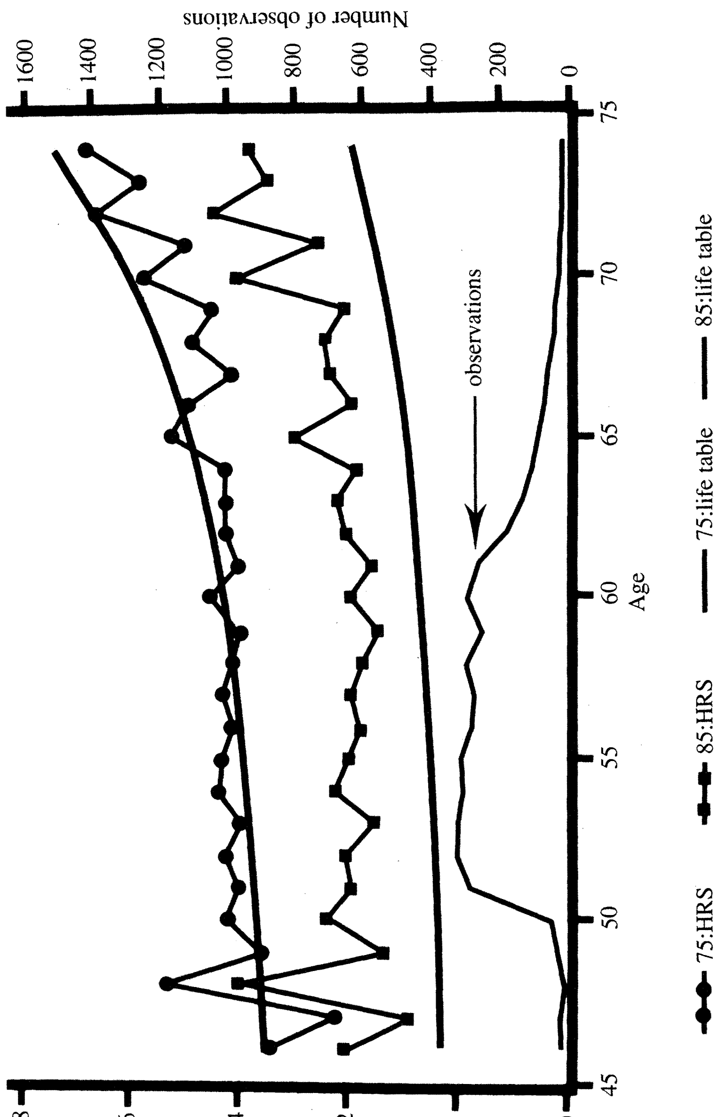

Figure 2
Probability of surviving: men

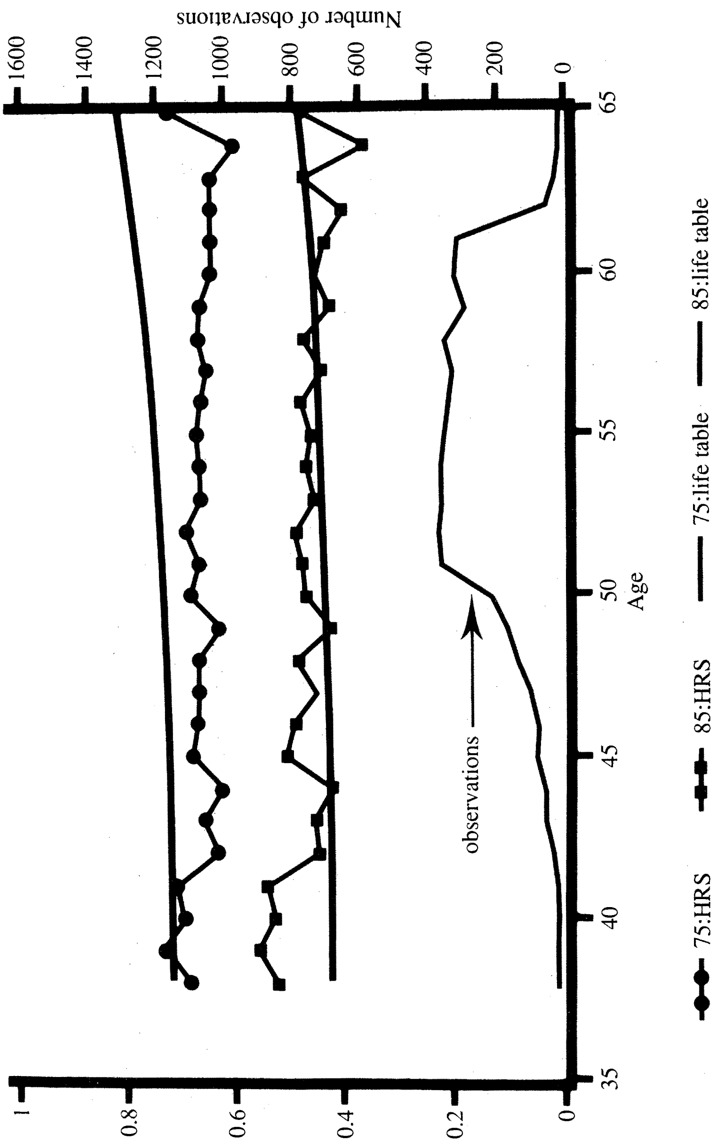

Figure 3
Probability of surviving: women

If mortality risk is stationary over time and there were no heterogeneity in the population, these paths should slope upward, reflecting a positive probability of dying in any year. However, neither of these conditions is met. As we have already seen in Table 1, reductions in age-specific mortality rates have increased the survival rates as measured in life tables. Therefore, a 50-year-old man could reasonably forecast future improvements in age-specific mortality risk such that his predicted survival probability to age 75 would be about the same as the survival probability of someone considerably older.[7] For example, if a 50-year-old man uses the year-2000 life table to form probabilities, his estimate of P75 would be about 0.60. If a 56-year-old man uses the 1990 life table his estimate would also be 0.60.

The effect of a shift in survival probabilities is even greater on P85. For example, based on a year-2000 life table the probability of a 50-year-old living to 85 is about 0.28. Based on the 1990 life table the probability of a 62-year-old man living to 85 is also about 0.28.

We do not know how people form their subjective probabilities about living to 75 or 85. But the rapid change in mortality risk leads us to conclude that a declining path of P75 and P85 with age, especially at younger ages, can be consistent with our thinking of them as probabilities.

B. Conditional Probabilities of Survival

At the individual level P85 should be less than P75, so that an individual's subjective probability of living to 85 conditioned on living to 75 is less than 1. Table 2 has examples of the distribution of P85 conditional on P75, and the average of P85. For example, 276 respondents gave P75 = 0.2. Among these respondents, the average of P85 is 0.055; 59 percent gave P85 = 0. The results shown for P75 = 0.2, 0.5, and 0.8 are typical: the average of P85 is considerably less than P75 at each of the nine points between 0 and 1.0; very little of the probability mass of P85 is greater than the conditioning value of P75 (shown by an underline in Table 2). The distribution of P85 given P75 = 1.0 is an exception: 41.8 percent of the respondents who gave P75 = 1.0 also gave P85 = 1.0. Some of this bunching could be due to the coarseness of the scale offered to respondents: a very optimistic person could round a high subjective probability to 1.0 when given the choice between 0.9 and 1.0. However, it seems unlikely that such a large fraction would be that optimistic. More likely the bunching reveals cognition error or misunderstanding, or somewhat less likely observation error. In any of these events, a researcher would have to model the process that leads to these responses.

Table 3 has information about the joint responses. About 70 percent of the individuals have P75 greater than P85. It is not clear how much cognition error or misunderstanding is revealed by the other 30 percent of the respondents: the

7. In addition to time trends in age-specific mortality risk, younger cohorts may have been exposed to fewer environmental insults and they may have led healthier lives. Thus, the cohort of 50-year-olds may be more healthy than the cohort of 62-year-olds were when the latter were 50, and could have higher survival probabilities than would be indicated by the trends in age-specific mortality risk.

Table 2
Means and Distribution of Probability of Living to 85 Conditional on Probability of Living to 75

Probability of Living to 85	Probability of Living to 75: 0.2	0.5	0.8	1.0
0.0	0.591	0.146	0.043	0.042
0.1	0.312	0.085	0.024	0.008
0.2	0.072	0.216	0.065	0.014
0.3	0.004	0.199	0.073	0.016
0.4	0.004	0.104	0.084	0.023
0.5	0.011	0.219	0.203	0.179
0.6	0.004	0.009	0.132	0.047
0.7	0	0.008	0.185	0.067
0.8	0	0.008	0.165	0.124
0.9	0	0.004	0.024	0.061
1.0	0.004	0.002	0.004	0.418
Average	0.055	0.283	0.451	0.749
Observations	276	1,710	786	1,746

Source: Authors' calculations from HRS early release.

Table 3
Comparison of Probabilities of Living to 75 and 85

Probability Comparison	Percent of Respondents
P75 > P85	70.1
Both probabilities = 0	6.9
Both probabilities = 0.5	4.7
Both probabilities = 1.0	9.2
Both probabilities = some other value	6.6
P75 < P85	2.5

Source: Authors' calculations from HRS.
Note: Based on 7,916 observations.

ties could be explained by uninformed guessing by the respondent or observation error, which would have to be modeled by an analyst. The respondents that gave zero for both probabilities could be pessimistic, and when required to choose between 0 and 1 on the scale of 0 to 10, chose zero. The 2.5 percent of respondents with P75 < P85 cannot be explained except as cognition or response error. The 9.2 percent with P75 = P85 = 1.0 may have been very optimistic and, when required to choose between 9 and 10 on the ten-point scale, chose 10. Perhaps more likely, they did not understand the nature of the question.

We thought that with age the responses might become more heterogeneous: as people get new information about their health status and as they age toward 75, they may either become convinced they will live to 75 or convinced they will not live. Thus, rather than the average coming from a population in which everyone had the same population probability, it would come from a population in which a fraction had probability one and another fraction probability zero so that the fractions averaged to the population probability. If the responses became more heterogeneous with age, the variance of P75 and P85 should increase with age. We studied the standard deviation of P75 and P85 as a function of age, but we could not find any pattern. Another indicator would be the fraction of zeros or ones, but again we found no trend in the fraction with age.

Even with the inconsistencies shown in Table 3, our overall assessment of P75 and P85 is quite positive. The response rate to the probability questions is very high, and it is likely that there is information even in the lowest quality responses. Broadly speaking the observations on P75 and P85 act like probabilities, and, given the changes in life tables over time, they aggregate to reasonable levels. Furthermore, all variables derived from household interviews have inconsistencies and observation errors, and analysts have econometric methods to handle these problems. We imagine that the inconsistencies in P75 and P85 will prove to be tolerable and that their inconsistencies and errors are probably no larger than those in many other variables.

C. *Covariation with Other Variables*

Even with changing mortality risk from cohort to cohort, at least the qualitative variation of P75 and P85 with risk factors should be consistent with epidemiological data if they are good predictors of mortality experience. For example, someone who smokes should have a lower probability of living to 75 than someone from the same cohort who does not smoke. Averaging the probabilities over smokers and nonsmokers from the same cohort will reveal such a difference. A difference will also be found after averaging over cohorts unless the incidence of smoking varies substantially with cohorts and there is a change in cohort-specific mortality risk. It is beyond the scope of this paper to study changes in risk factors by cohort, so we will assume that the incidence of risk factors is roughly constant.

We will find if P75 and P85 vary qualitatively with risk factors as they do in epidemiological data. However, just as in the epidemiological data that we use for comparisons, we do not particularly ascribe causality to the relationships. Rather our objective is to find if the variation with risk factors in anticipated survival in the HRS population is the same as it is in actual mortality data.

1. Variation with Socioeconomic Status

It is well known that mortality risk varies with a number of indicators of socioeconomic status: education, wealth, and income to name but a few (Kitagawa and Hauser 1973; Silver 1972; Berkman 1988; Feinstein 1993). Table 4 shows that those in the highest income quartile give average survival probabilities to age 75 that are 0.11 higher than those in the lowest quartile. We estimate that this is a difference in life expectancy of three to four years (Hurd and McGarry 1993a). The variation by wealth quartile is about the same as the variation by income quartile.

Table 5 shows that the survival probabilities increase with educational level as they do in epidemiological data. The variation is comparable to the variation by income or wealth quartiles.

2. Variation with Health Status and Risk Factors

The HRS respondents were asked to assess their own health. Table 6 has the distribution of responses to the health question and the averages of P75 and P85 by health status. Most respondents rate their health as good, very good, or excellent. The variation in the survival probabilities is enormous: P75 ranges from 0.34 to 0.75 among men and 0.40 to 0.78 among women, and the variation in P85 is similar.

Within health categories, P75 and P85 are higher among women than among men. Women have fewer risk factors such as smoking and heavy drinking. For example, in our data 54 percent of women are current or former smokers, compared with 74 percent of men; 2 percent of women drink three or more drinks per day, compared with 9 percent of men. Further, there are surely other unobserved determinants of longevity that vary by sex, even holding health status constant.

Tables 7 and 8 show the probabilities of survival as a function of smoking and drinking. Smoking, of course, is a risk factor in the population (Lew and Garfinkel 1987), and corresponding variations in P75 and P85 are found in Table 7. Furthermore, the difference between "never smoked" and "not now" (but in the past) is rather small, just as it is in the population at large (Rogers and Powell-Griner 1991). The mortality risk of current smokers relative to those who have never smoked is 1.21.[8] Thus, the probability of smokers dying before age 75 is 21 percent greater than the probability of never-smokers dying before age 75. This is somewhat less than the relative risk across income quartiles, which is 1.37 for the lowest income quartile relative to the highest.

In epidemiological data, moderate drinking is associated with greater longevity than complete abstinence, and heavy drinking (five or more drinks per day), with substantially reduced longevity (Shaper 1990; Boffetta and Garfinkel 1990; Ellison

8. Relative mortality risk is the probability of dying among those in some risk group divided by the probability of dying among those in a reference group. It is widely used in epidemiological studies because it normalizes on a baseline probability, and so adjusts for scaling. Our calculation of relative risk is Q_{rg}/Q_r where the Q's are, respectively, 1 − P75 of a risk group (say, current smokers) and 1 − P75 of the reference group (say, never-smokers).

Table 4
Average Probability of Living to 75 or 85: Income and Wealth

	To 75		To 85	
Quartile	Income	Wealth	Income	Wealth
Lowest	0.59	0.57	0.39	0.39
Second	0.63	0.62	0.42	0.40
Third	0.66	0.66	0.43	0.44
Highest	0.70	0.70	0.48	0.47

Source: Authors' calculations from HRS.
Note: Based on 7,900 observations.

Table 5
Average Probability of Living to 75 or 85: Education

Education Level	Observations	To 75	To 85
Less than high school	2,190	0.57	0.37
High School	2,855	0.65	0.42
Greater than high school	2,896	0.69	0.48

Source: Authors' calculations from HRS.

Table 6
Average Probability of Living to 75 or 85: Self-Assessed Health Status

	Men			Women		
Health Status	Observations	To 75	To 85	Observations	To 75	To 85
Excellent	793	0.75	0.53	1,006	0.78	0.58
Very good	998	0.68	0.42	1,236	0.71	0.50
Good	1,037	0.61	0.37	1,162	0.64	0.44
Fair	449	0.47	0.27	645	0.53	0.33
Poor	286	0.34	0.16	328	0.40	0.23

Source: Authors' calculations from HRS.

Table 7
Average Probability of Living to 75 or 85: Smoking Status

Smoking Status	Observations	To 75	To 85
Never smoked	2,927	0.67	0.47
Formerly but not now	2,878	0.65	0.43
Yes	2,138	0.60	0.38

Source: Authors' calculations from HRS.

Table 8
Average Probability of Living to 75 or 85: Drinking

Drinks per Day	Observations	To 75	To 85
None	3,126	0.61	0.41
Fewer than 1	3,593	0.67	0.45
1–2	812	0.68	0.44
3–4	295	0.60	0.36
5 or more	112	0.55	0.33

Source: Authors' calculations from HRS.

1990). This is precisely what is found in the variation in P75 and P85 by the amount of drinking.[9]

Table 9 has the averages of the survival probabilities by health and education. Because of the positive correlation between health and education the effects of education on the survival probabilities are much smaller than when health is not kept constant. For example, at health levels of good or very good (where most of the observations lie), P75 varies by just 0.03 or 0.04 with educational level, and P85 by even less. Even so, the probabilities mostly increase with educational level, particularly for education beyond high school. Within educational levels the survival probabilities vary with health status by about as much as they do in Table 6. These results show that education as a measure of socioeconomic status does not account for a good deal of the information about survival that individuals have and which they reveal in the subjective probabilities and self-assessed health.

We obtain similar results when we interact health status with smoking status, or with income or wealth quartiles (not shown). Within health categories the

9. Again, we do not mean to imply causality, which, in the case of alcohol consumption, is the subject of considerable controversy (Shaper 1990).

Table 9
Average Probability of Living to 75 or 85: Health Status and Education

Health Status	Education: Less than High School	High School	More than High School
	Living to 75		
Excellent	0.71	0.77	0.78
Very good	0.68	0.69	0.71
Good	0.60	0.63	0.64
Fair	0.51	0.48	0.52
Poor	0.36	0.34	0.44
	Living to 85		
Excellent	0.52	0.53	0.58
Very good	0.45	0.45	0.48
Good	0.40	0.40	0.41
Fair	0.33	0.26	0.32
Poor	0.19	0.17	0.25

Source: Authors' calculations from HRS.
Note: Based on 7,935 observations.

variation in the probability of survival is much smaller than in Tables 4, 7, and 8; yet, among smokers or within an income or wealth quartile, the probabilities vary substantially with health status. Apparently the main result of smoking is to affect self-assessed health (and actual health), which, in turn, changes life expectancy. The main effect of income or wealth is to signal differences in health status.

When we interact health with drinking status, moderate drinking (holding health constant) increases the survival probabilities only slightly over abstinence. Yet, the difference between heavy drinking and moderate drinking is about as large as in Table 8, which has no control for health status. For example, when health status is very good, those who drink 1–2 drinks per day reported an average P75 of 0.71, whereas heavy drinkers (5+) reported an average of 0.59.

Subjective probability of surviving to 75 among those whose health status is very good, by number of drinks per day

Drinks	None	less than 1	1–2	3–4	5 or more
P75	0.69	0.70	0.71	0.65	0.59

The comparable numbers from Table 8 are 0.68 and 0.55, respectively. Perhaps heavy drinkers whose health is very good anticipate a decline in health status and they have incorporated that into their subjective probability distributions.

The variation in P75 and P85 by health and drinking status is remarkably large (not shown). For example, P75 is 0.74 among those who do not drink at all and are in excellent health; it is 0.30 among those who are heavy drinkers and in poor health. Because of such large variation in P75 and P85, the accuracy of their predictions of the future mortality experience of the HRS sample can be checked in panel even at the relatively young ages of the HRS respondents, whose average mortality risk is low.

3. Mortality Experience of Parents

In the population, the longevity of children increases with the longevity of their parents (Feinstein 1993). We imagine, however, that the functional relationship between the mortality experience of the parents and the child's mortality risk is rather complicated. For example, in that the leading cause of death at an early age is accidents, the effects of the very early death of a parent on P75 or P85 will probably be qualitatively different from the effects of death at a later age. In particular, the effect will not be monotonic with the age of the parents' death.

Table 10 shows the average probabilities of survival by the mortality experience of the parents of the HRS respondents. The variation is what we would expect: if the parents are alive, the probabilities are higher than if the parents have died; if the parents died before age 51, the probabilities are higher than if they died between 51 and 64; at later ages the survival probabilities increase monotonically with the parents' ages at death. The results hold separately for both the mother and father and for survivorship to age 85. The standard errors are small, and in many cases the differences are significant.

A close examination of the difference between P75 and P85 shows that the effect of increasing the age of a parent's death from 65–74 to 75–84 is greater on P75 than on P85, compared with the effect of increasing the age from 75–84 to 85+. This can be shown more clearly by the variation in relative risk, which is the probability of dying among those in some risk group relative to the probability of dying among those in a reference group. In each of our calculations, the risk group will be those whose mothers died at older ages and the reference group will be those whose mothers died at younger ages. Thus, in Table 11, the first entry gives the average subjective probability of dying before 75 among those whose mother died between the ages of 51–64 relative to the average subjective probability of dying among those whose mother died before 51. The relative risk is 1.056; the mother's dying between 51 and 64 rather than before 51 increases the respondent's subjective probability of dying before 75 by 5.6 percent. The second entry (0.948) gives the risk of those whose mother died between the ages of 65–74 relative to those whose mother died between the ages of 51–64. Relative risk being less than 1 derives from the increase in P75 with the age of the mother at her death.

The relative risk of dying before 75 is smallest when the mother died between 75 and 84, whereas the relative risk of dying before 85 is smallest when the mother

Table 10
Average Probability of Living to 75 or 85, by Survivorship of Parents

		To 75		To 85	
Survivorship of Parents	Observations	Mean	Standard Error	Mean	Standard Error
Mother alive	3,424	0.689	0.004	0.480	0.005
Mother's age at death					
<51	565	0.595	0.012	0.377	0.012
51–64	850	0.572	0.009	0.360	0.010
65–74	1,113	0.594	0.008	0.364	0.008
75–84	1,197	0.644	0.008	0.407	0.008
85+	462	0.676	0.012	0.478	0.013
Age missing	195	0.607	0.020	0.398	0.021
Father alive	1,382	0.693	0.006	0.495	0.007
Father's age at death					
<51	725	0.624	0.010	0.409	0.010
51–64	1,407	0.604	0.007	0.380	0.007
65–74	1,776	0.619	0.006	0.394	0.007
75–84	1,508	0.677	0.006	0.445	0.007
85+	518	0.714	0.011	0.516	0.012
Age missing	399	0.619	0.013	0.426	0.014

Source: Authors' calculations from HRS.

died at 85 or over. Apparently special importance is placed on the mother's survival to 75 in the respondent's forming a subjective probability of surviving to 75, but it is less important that the mother survived well past the age of 75. Similarly, in forming P85, particular importance is put on the mother's survival past 85, and less on whether she survived beyond the age interval 65–74 to the age interval 75–84. This relationship holds for both the mortality experience of the mother and of the father. These results imply that respondents are sensitive to the parents' mortality experience and adjust their probabilities of survival to that experience.

We conclude from the cross-tabulations by socioeconomic status, health status, risk factors, and parents' mortality experience that the variation in P75 and P85 is qualitatively in agreement with epidemiological data along all the dimensions we have discussed. We have noted particularly the influence of the parents' mortality experience, which even extends to different effects on P75 and P85.

D. Regression

Further cross-tabulations that control simultaneously for a number of risk factors are not practical, so we estimated linear regressions of P75 and P85 on measures of socioeconomic status, personal characteristics, risk factors, and diseases.

Table 11
Relative Risk of Dying before 75 or 85, by Age of Parents at Death

Risk Group	Reference Group	Before 75 Mean	Before 75 Standard Error	Before 85 Mean	Before 85 Standard Error
Mother's age at death					
51–64	50 or less	1.056	0.026	1.027	0.041
65–74	51–64	0.948	0.021	0.993	0.035
75–84	65–74	0.878	0.018	0.934	0.030
85 or over	75–84	0.911	0.021	0.880	0.034
Father's age at death					
51–64	50 or less	1.053	0.020	1.048	0.032
65–74	51–64	0.962	0.016	0.979	0.026
75–84	65–74	0.848	0.014	0.915	0.024
85 or over	75–84	0.884	0.018	0.873	0.029

Source: Authors' calculations from Table 10.
Note: Relative risk is the probability of dying of a risk group divided by the probability of dying of a reference group. The risk groups and reference groups are defined by the age of the mother or father at death. Standard errors are asymptotic standard errors.

Table 12 has estimated coefficients and standard errors for P75 and P85. Income has a small, not significant coefficient; wealth has a small coefficient with a t-statistic just over 1.95. We say these are small in that the variation in P75 explained by the coefficients on income and wealth as income and wealth vary across quartiles is small compared with the actual variation across the quartiles (Table 4).

The change with age is much smaller than what is found in a life table: here ten years change P75 by about 0.04 compared with about 0.15 in the 1988 life table. We have already discussed how cohort effects could account for the difference.

Although there is some variation with the amount of physical activity, apparently the measures of physical activity, both light and heavy, mainly classify people into those who are physically active and those who are not (never exercise). Not being physically active reduces P75 by 0.05 to 0.06, and it matters little if the physical activity is light or heavy. The result, of course, does not imply that exercise will increase longevity because health status will influence both whether people are physically active and longevity. In particular, many who are never physically active are likely to be incapable of physical activity.

The regression implies that, ceterus paribus, whites report lower subjective probabilities of survival than blacks. Yet, in life tables elderly whites have greater life expectancies than elderly blacks.[10] A possible explanation is that the regres-

10. For example, life expectancy of 65-year-old males in 1990 was 15.2 among whites and 13.2 among blacks, and the difference was similar for women (U.S. Bureau of the Census 1993).

Table 12
Determinants of Probability of Living to 75 or 85

Variable	To 75 Parameter	To 75 Standard Error	To 85 Parameter	To 85 Standard Error
Intercept	0.764	0.022	0.604	0.024
Household income (100 thousand)	0.010	0.009	0.008	0.009
Wealth (millions)	0.016	0.008	0.007	0.009
Age	0.004	0.001	0.003	0.001
Married	−0.001	0.008	−0.007	0.009
Male	−0.047	0.008	−0.074	0.009
Light physical activity: 3+ per week	—	—	—	—
1–2 per week	−0.018	0.008	−0.026	0.009
1–3 per month	−0.017	0.012	−0.025	0.013
<1 per month	−0.024	0.013	−0.016	0.014
Never	−0.054	0.012	−0.052	0.013
Heavy physical activity: 3+ per week	—	—	—	—
1–2 per week	−0.008	0.013	−0.009	0.015
1–3 per month	−0.009	0.014	−0.020	0.016
<1 per month	−0.043	0.012	−0.040	0.013
Never	−0.062	0.011	−0.067	0.012
Race (white = 1)	−0.049	0.008	−0.091	0.009
Formerly smoked	0.000	0.008	−0.010	0.009
Currently smokes	−0.033	0.009	−0.046	0.009
Does not drink	—	—	—	—
Drinks <1 per day	0.022	0.007	0.012	0.008
Drinks 1–2	0.026	0.012	0.005	0.013
Drinks 3–4	−0.014	0.018	−0.023	0.019
Drinks 5+	−0.027	0.028	−0.024	0.030
Education <12	−0.041	0.008	−0.015	0.009
Education >12	0.023	0.008	0.042	0.009
Ever high blood pressure	−0.043	0.007	−0.047	0.008
Ever diabetes/high blood sugar	−0.054	0.011	−0.060	0.012
Cancer/malignant tumor	−0.062	0.014	−0.033	0.016
Chronic lung disease	−0.067	0.012	−0.068	0.013
Ever heart problems	−0.067	0.012	−0.091	0.013
Angina/chest pains	−0.058	0.019	−0.030	0.021
Congestive heart failure	−0.058	0.026	0.013	0.029
Ever had stroke	−0.030	0.020	0.008	0.022
Arthritis/rheumatism	−0.022	0.007	−0.030	0.008
Weight (100 pounds)	0.001	0.001	0.002	0.001

Source: Authors' calculations from HRS.
Note: Average of P75 = 0.645 based on 7,741 observations; R^2 = 0.102. Average of P85 = 0.433 based on 7,712 observations; R^2 = 0.093.

sion coefficient results from standardizing on all the other variables in the regression. This explanation is inadequate, however, because the results are similar in simple cross-tabulations:

	White	Black
P75	0.642	0.654
P85	0.419	0.481

Thus, there is no obvious explanation for the difference in the subjective probabilities.

Smoking is associated with a significant difference in survival probabilities, but the effect is smaller than in the cross-tabulations: 0.03 compared with 0.07. Of course, the regression controls for risk factors that vary with smoking status such as education, income, and health conditions.

Similarly the effects of drinking and education are the same qualitatively as in the cross-tabulations, but the effects are attenuated.

The incidence or prevalence of diseases affects P75 as would be expected: all are negative, and many of them are large, reducing survival probabilities by 0.06 to 0.07. For example, having had cancer or a malignant tumor reduces P75 by .062. We estimate this to be a reduction in life expectancy of two to three years. To the extent that smoking affects life expectancy by causing these diseases, including them in the regression will attenuate the effects of smoking, which is what we observe.

There are some differences in the effects of diseases on P75 and P85, but roughly speaking they are comparable.

In additional results not shown we added the self-assessed health measures to the regressions. The health variables have large coefficients: for example, the variation in P75 is -0.34 as health varies from excellent to poor, which is not much less than in the cross-tabulations of Table 6. Adding the health variables increased the R^2 substantially: from 0.10 to 0.16 for P75 and 0.09 to 0.13 for P85. The effects of other variables on P75 and P85 are attenuated but not eliminated by including the health variables. Both the probabilities and self-assessed health are subjective measures based partly on how the respondent feels. It is clear that they are related to each other, but they are not the same: other variables given health status have the expected qualitative effects.

We found in cross-tabulations that the mortality experience of the parents has an important and predictable relationship with P75 and P85. When variables describing the mortality experience of parents are added to the regressions they have important magnitudes, much like what is found in the cross-tabulations.

We extended these results based on the notion that because men and women die from different causes, men may be more influenced by the father's mortality experience and women more by their mother's experience. We tested this by regressions of the type reported in Table 12, but with the addition of 16 variables describing the mortality experience of the parents. We estimated the regressions separately by sex of the respondent.

Table 13
Effect of Parents' Age or Age of Death on Probability of Living to 75 or 85, by Sex of Respondent

	Living to 75		Living to 85	
	Females	Males	Females	Males
Parent alive				
Mother (1 = yes)	0.072[a]	0.077[a]	0.073[a]	0.083[a]
Father (1 = yes)	0.055	0.071	0.077[a]	0.084
Mother's age −65	0.0021	0.0006	0.0058[a]	0.0007
Father's age −65	−0.0004	0.0014	0.0012	0.0021
Age of parent at death				
Mother's age <51 (1 = yes)	0.034	0.053[a]	0.074[a]	0.031
Father's age <51 (1 = yes)	0.001	0.050[a]	0.023	0.051[a]
50 < mother's age < 65 (1 = yes)	−0.008	0.025	0.027	0.045
50 < father's age < 65 (1 = yes)	−0.003	0.001	0.015	0.014
Mother's age −65	0.0050[a]	0.0024[a]	0.0072[a]	0.0034[a]
Father's age −65	0.0029[a]	0.0055[a]	0.0047[a]	0.0061[a]

Source: Authors' calculations from HRS.
Note: Extract from regressions with 47 right-hand-side variables, based on 4,712 observations on females and 3,365 observations on males (P75). $R^2 = 0.126$ for females and 0.142 for males (P75).
a. Significant at 5 percent level (2-tailed test).

An extract of these regressions is in Table 13. The reference group is those whose parents died at age 65. The regressions have two types of variables to describe the parents' mortality experience: categorical variables, which are indicated in the table by (1 = yes), and variables that are continuous in either the parents' age or age of death. The ages are normalized to be zero at 65.

Among female respondents, if the mother is alive, P75 is predicted to be 0.072 greater (the coefficient on the "mother" categorical variable) than if the mother had died at 65, and it increases by 0.0021 for each year of the mother's age (coefficient on "mother's age 65"). If the father is alive, P75 is greater by 0.055, and it decreases by 0.0004 per year of age. If the mother died before age 51, P75 is higher by 0.034 than if she had died at age 65. If the mother died between 51 and 64, P75 is almost the same as if her death had been at age 65. If the mother died at an age greater than 65, her age at death increases P75 by 0.0050 per year, so that if she died at 85, P75 would be higher by 0.10 than if she had died at age 65. This is, of course, a rather large difference in P75.

The effects on P85 are similar.

It seems clear from these results that the respondents are aware that the age of their parents or the age of their parents' death has an influence on their own mortality risk and that they consistently alter their reports on P75 and P85. The effects are large, particularly because the regressions control for other risk factors

Table 14
Fitted Probability of Survival to 75: Effects of Survivorship of Parents

Survivorship of Parent, and Age or Age at Death	Female Respondents		Male Respondents	
	Mother	Father	Mother	Father
Alive				
65	0.69	0.68	0.72	0.71
85	0.74	0.67	0.73	0.74
Deceased				
65	0.62	0.62	0.64	0.64
85	0.72	0.68	0.69	0.75

Source: Calculated from Table 13.

which are probably associated with the lifetime health status of the parents and their age at death.

What is most striking about Table 13, however, is the considerable symmetry, particularly among the coefficients on the ages at which the parents died. For example, the quantitatively important coefficients on mother's age at death and father's age at death are roughly twice as large on the same-sex parent as on the opposite-sex parent.

These results are illustrated in Table 14, where we have calculated from the regression coefficients the fitted value of the survival probability to age 75 as a function of the mortality experience of the parents of the respondents. The table shows that P75 is considerably greater if parents (either mother or father) are alive and are age 65 than if they had died at 65 (the first row compared with the third row). If the parents died at 85 the subjective probabilities are about the same as if the parents are still alive and are 85 (second row compared with the last row). This implies it is important that parents survive well past 75, but not particularly important that they survive into their 90s. The increase in P75 from having the same-sex parent survive is much greater than from having the opposite-sex parent survive. For example, among female respondents P75 increases by 0.10 if the mother died at 85 rather than at 65, whereas the increase is 0.06 if the father died at 85 rather than at 65. On average P75 increases by 0.07 when the same-sex parent survives to 85 but by just 0.02 when the opposite-sex parent survives to 85.

IV. Conclusion

Our criteria for judging the measures of the subjective probabilities of survival in the HRS were that they are good approximations to population probabilities, that they are internally consistent, and that they covary with other variables in the same way as in other data. On average the probabilities of living to 75 or 85 are close to averages in a life table from 1990. However, in view of the rapid change in mortality rates, it is not really clear how close they should be because we do not know how cohorts form their views about mortality risk.

The subjective probabilities are in general internally consistent. To the extent that they are not, an analyst should model the process that causes the inconsistency. The process surely includes observation error, and in this regard subjective probabilities are no different from almost all economic variables. Usually, however, the respondent and the analyst share a common understanding of the meaning of a survey question. This is undoubtedly not always true for the questions about subjective probabilities, and that difference needs to be taken into account.

The sharpest test of the subjective probabilities comes from their covariation with other variables. The probabilities of living to 75 or 85 vary in a systematic and reasonable way with diseases, socioeconomic status, self-assessed health, and indicators of family longevity. Of course, what we hope is that conditional on observable characteristics, the subjective probabilities will be good predictors of longevity that will allow us to observe and control for individual heterogeneity. Finding whether this happens will require observations in the panel data. From the cross-section, however, we conclude that the measures of subjective probabilities in the HRS show great promise of making a substantial contribution to our understanding of decision-making under uncertainty.

References

Anderson, Kathryn H., Richard V. Burkhauser, and Joseph F. Quinn. 1986. "Do Retirement Dreams Come True? The Effect of Unexpected Events on Retirement Age." *Industrial and Labor Relations Review* 39(4):518–26.

Berkman, Lisa F. 1988. "The Changing and Heterogeneous Nature of Aging and Longevity: A Social and Biomedical Perspective." In *Annual Review of Gerontology and Geriatrics, 8,* eds. G. L. Maddox and M. P. Lawton, 37–68.

Bernheim, B. Douglas. 1988. "Social Security Benefits: An Empirical Study of Expectations and Realizations." In *Issues in Contemporary Retirement,* eds. E. Lazear and R. Ricardo-Campbell, 312–48. Stanford: Hoover Institution.

———. 1989. "The Timing of Retirement: A Comparison of Expectations and Realizations." In *The Economics of Aging,* ed. D. Wise 335–55. Chicago: University of Chicago Press.

———. 1990. "How Do the Elderly Form Expectations: An Analysis of Responses to New Information." In *Issues in the Economics of Aging,* ed. D. Wise, 259–83. Chicago: University of Chicago Press.

Boffetta, Paolo, and Lawrence Garfinkel. 1990. "Alcohol Drinking and Mortality among Men Enrolled in an American Cancer Society Prospective Study." *Epidemiology* 1(5):343–48.

Dominitz, Jeff. 1993. "Subjective Expectations of Unemployment, Earnings, and Income." University of Wisconsin–Madison, typescript.

Dominitz, Jeff, and Charles Manski. 1994a. "Using Expectations Data to Study Subjective Income Expectations." NBER Working Paper 4937, Cambridge, Mass.

———. 1994b. "Eliciting Student Expectations of the Returns to Schooling." Department of Economics, University of Wisconsin–Madison, typescript.

Ellison, R. Curtis. 1990. "Cheers!" *Epidemiology* 1(5):337–39.

Feinstein, J. 1993. "The Relationship between Socioeconomic Status and Health: A Review of the Literature." *The Milbank Quarterly* 71(2):279–322.

Hall, Arden, and Terry Johnson. 1980. "The Determinants of Planned Retirement Age." *Industrial and Labor Relations Review* 33:240–55.

Hamermesh, Daniel S. 1985. "Expectations, Life Expectancy, and Economic Behavior." *Quarterly Journal of Economics* 100(2):389–408.

Hamermesh, Daniel S., and Frances W. Hamermesh. 1983. "Does Perception of Life Expectancy Reflect Health Knowledge?" *American Journal of Public Health* 73(8):911–14.

Hurd, Michael D. 1989. "Mortality Risk and Bequests." *Econometrica* 57(4):779–813.

Hurd, Michael D., and Kathleen McGarry. 1993a. "Evaluation of Subjective Probability Distributions in the HRS." NBER Working Paper 4560, Cambridge, Mass.

———. 1993b. "The Relationship between Job Characteristics and Retirement." NBER Working Paper 4558, Cambridge, Mass.

Juster, F. Thomas, and Richard Suzman. 1993. "The Health and Retirement Study: An Overview." Presented at the HRS Early Results Workshop, University of Michigan, September.

Kitagawa, Evelyn M., and Philip M. Hauser. 1973. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Cambridge, Mass.: Harvard University Press.

Lew, Edward A., and Lawrence Garfinkel. 1987. "Differences in Mortality and Longevity by Sex, Smoking Habits, and Health Status." *Transactions: Society of Actuaries* 39:107–30.

Rogers, Richard G., and Eve Powell-Griner. 1991. "Life Expectancies of Cigarette Smokers and Nonsmokers in the United States." *Social Sciences and Medicine* 32(10):1151–59.

Shaper, A. G. 1990. "Alcohol and Mortality: A Review of Prospective Studies." *British Journal of Addiction* 85:837–47.

Silver, Morris. 1972. "An Econometric Analysis of Spatial Variations in Mortality Rates by Age and Sex." In *Essays in the Economics of Health and Medical Care*, ed. Victor Fuchs, 161–227. New York: Columbia University Press.

U.S. Bureau of the Census. 1993. "Statistical Abstract of the United States, 1993." Washington, D.C.: GPO.