THE PREDICTIVE VALIDITY OF SUBJECTIVE PROBABILITIES OF SURVIVAL

Michael D. Hurd
Kathleen McGarry

Working Paper 6193

NBER WORKING PAPER SERIES

THE PREDICTIVE VALIDITY OF SUBJECTIVE PROBABILITIES OF SURVIVAL

Michael D. Hurd
Kathleen McGarry

Working Paper 6193
http://www.nber.org/papers/w6193

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1997

Hurd thanks the National Institute on Aging for financial support through a grant to the National Bureau of Economic Research. McGarry gratefully acknowledges financial support from the Brookdale Foundation. This paper is part of NBER's research program in Aging. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

The Predictive Validity of Subjective
Probabilities of Survival
Michael D. Hurd and Kathleen McGarry
NBER Working Paper No. 6193
September 1997
JEL Nos. I12, J14, D84
Aging

ABSTRACT

Although expectations (or subjective probability distributions) play a prominent role in models of decision-making under uncertainty, we have very little data on them, and are instead forced to base our models on unverifiable assumptions. Macroeconomic models often assume rational expectations, and microeconomic models base estimation on observable population probabilities. An alternative to these assumptions is to query individuals directly about their subjective probabilities, and to use the responses as measures of expectations. Prior research on subjective survival probabilities in the Health and Retirement Study has shown that reported probabilities aggregate closely to life table values and covary appropriately with known risk factors. This paper uses panel data to study the evolution of subjective survival probabilities and their ability to predict actual mortality. We find that respondents modify appropriately their survival probabilities based on new information. The onset of a new disease condition or the death of a parent between the waves is associated with a reduction in survival probabilities. The subjective survival probabilities also predict actual survival. Those who survived in our panel reported probabilities approximately 50 percent greater at baseline than those who died. Although more needs to be learned about properties of subjective probabilities, we conclude that they show considerable promise for estimating models of decision-making under uncertainty.

Michael D. Hurd
RAND Corporation
1700 Main Street
Santa Monica, CA 90407
and NBER
mhurd@rand.org

Kathleen McGarry
Department of Economics
University of California, Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90095-1477
and NBER
mcgarry@ucla.edu

1 Introduction

Expectations play a prominent role in many kinds of economic models because agents are thought to make decisions in the present that are at least partly based on what they believe will happen in the future. Although we use the word "expectations," it is the probability distributions of future events that typically enter economic models. For example, it is not life expectancy that helps determine consumption in a life cycle model, but rather the probability distribution of the random date of death.[1] In fact, it is not even the objective probability, but rather the individual's subjective belief that determines behavior.

Despite their importance, we have had very little data on subjective probabilities, so it has been necessary to make assumptions about them. For example, in macroeconomic models it is common to assume that expectations are rational. The mechanism for achieving rational expectations is that an agent with views that are persistently incorrect (or one that uses less than all the available information) will be driven from business or be forced into better behavior. This is a rather strong assumption and it is untestable without data on actual expectations. Furthermore, while the motivating story is plausible for some kinds of agents (such as bond traders) it is not plausible for others. An individual who persistently makes bad investment decisions about his personal savings will live to make still further bad decisions in the future.

In micro estimation it is often specified that individuals base their behavior on observable population probabilities. For example, estimation of life cycle consumption models has been based on the assumption that individuals approximately determine their behavior from the survival probabilities found in a life table. There are at least two objections to this assumption. The first is that the life tables may not be correct even on average. A population may anticipate gains in mortality that are not found in a life table, even a cohort life table. The second objection is that individuals in a population undoubtedly have differing subjective probability distributions and we can understand better individual behavior if we have information about the individual-level distributions.

An improvement to using population probabilities in microeconomic models is to base estimation on individuals' subjective probabilities. The subjective probabilities would be constructed from responses in household surveys about the probabilities of important events such as survival, health status, and retirement. For example, in life cycle models of consumption behavior, the level and rate of change of consumption depends on the individual's perceived survival curve, which can be estimated from life tables

[1] For simplicity we will refer to subjective probabilities as expectations as long as there no ambiguity.

and subjective survival to some particular age (Hurd, McFadden and Gan, 1995). Before we can confidently use individual-level probability distributions, however, we need to understand their properties. We would like to know whether individuals understand questions about probabilities and accurately report their beliefs about the likelihood of future events, whether individuals adjust their reported probabilities in response to new information, and whether these reported probabilities predict outcomes. If we can establish these properties, we will be more confident that the subjective probabilities provide information that is relevant to the individual's decision-making process, and therefore information that will be of use in our economic models.[2]

There has been little research on the determinants and predictive power of subjective probabilities. Juster (1966) tested whether the probabilities of an automobile purchase predicted actual purchases, and concluded that they did. Nonetheless, as discussed by Manski (1990), subjective probabilities have not been viewed favorably by economists and have not systematically been used in economic models. Rather data on "intentions" have been used. Manski has argued that "intentions" data are of limited usefulness and that subjective probabilities are superior for predicting outcomes.[3] Recently there has been a renewed interest in subjective probabilities such as individual probabilities of future earnings (Dominitz and Manski, 1994; Dominitz, 1997).

It is reasonable to suppose that individuals have expectations that are sufficiently well formed to be able to answer questions about subjective probabilities. When individuals choose consumption, savings, hours of work, or years of employment to maximize expected lifetime utility, they form expectations about future events. Their eventual utility depends on the accuracy of their forecasts, particularly an assessment of life expectancy: if someone projects too short a life span, he will likely save an insufficient amount to finance his retirement years and face unpleasantly low consumption levels later in life. Conversely, an individual who projects too long a life span will die with unspent wealth and, absent a bequest motive, with lower lifetime utility than if he had consumed the wealth.

Analysis of individual reports of survival probabilities has been limited. Hamermesh and Hamermesh (1983) and Hamermesh (1985) used non-population representative samples (including a sample of Ph.D. economists) to study individual reports of life expectancy and survival probabilities.

[2]Even if subjective probabilities have these properties, we will not really know of their usefulness in economic models until we find if they help explain behavior. We do not yet have the data to make such an assessment, but eventually the household surveys in progress will provide them.

[3]"Intentions" are represented by answers to questions such as "Do you intend to purchase an automobile?" and if so "Would it be certain, very likely, or likely?"

They found that expected length of life varied appropriately with factors such as smoking behavior, and obesity, although differences between groups were, for the most part, larger than actuarial differences. We know of no studies that have compared individual reports of subjective probabilities of survival with actual mortality experience.

The Health and Retirement Study (HRS) has fielded questions about expectations among a population-representative sample. The HRS elicited information about the respondents' subjective probability distributions including questions about the respondent's chances of surviving to age 75 and to age 85.[4] Response to these questions from the baseline interview appear to give data that can reasonably be interpreted to be subjective survival probabilities (Hurd and McGarry, 1995): their averages are close to averages from life tables, and they covary with know risk factors in an appropriate way. For example, men give lower probabilities than women, smokers give lower probabilities than nonsmokers, and those in higher socio-economic classes give higher probabilities of survival. Therefore, on average, the subjective survival probabilities will predict, at least qualitatively, actual mortality outcomes.

These results are based on cross-section analysis, so we do not know if individual reports of subjective probabilities provide information about mortality beyond what is already known about differential mortality across identifiable groups. In particular, we do not know whether, holding known risk factors constant, the subjective survival probabilities have any additional power to predict mortality. Furthermore, although we can study their determinants in cross-section, we would like to know how they evolve as new information comes to the respondent.

The broad goal of this paper is to give results that will increase our confidence that subjective probabilities can be used in models of decision-making under uncertainty: By stating a subjective probability has a respondent conveyed information about the probability distribution that is used in decision making? We will assess the reports of subjective survival probabilities from the second wave of HRS to see if they show the same broad patterns as in wave 1. We ask whether respondents give similar answers on average to a question about something that should be fairly stable over a two-year period. We will find the determinants of a change in the survival probabilities from wave 1 to wave 2 at the individual level. We ask if the probabilities change in a systematic and reasonable way with the arrival of new information. We will find if the reported survival probabilities in wave 1 help predict observed mortality outcomes between the waves. By controlling for observable risk factors such as

[4]Other questions included the chances of double-digit inflation and depression, the chances of working full-time past age 62 or 65, of moving and entering a nursing home.

smoking and disease conditions we can find if the respondent can give us information about his survival chances that would otherwise not be observable.

2 Data

The Health and Retirement Study (HRS) is a biennial panel survey of individuals born in the years 1931 through 1941 and their spouses. In 1992 when the first round of interviews was conducted, the sample was representative of the community-based U.S. population aged approximately 51-61.[5] The baseline sample contains 12,652 observations. The second wave of data from the HRS was collected in 1994, and 11,492 of the original 12,652 respondents were interviewed.[6] Most of our analysis will be based on the subsample who were 46-65 at the first interview. We use this sample rather than the age-representative sample in order to increase the number of observations, and in particular, the number of observed deaths. In addition we exclude wave 1 proxy interviews because the subjective probability questions were not asked of proxy respondents. With these restrictions our sample consists of 11,090 individuals in the first wave.

In each wave the HRS collected extensive information about health, cognition, economic status, work, and family relationships.[7] The observation on the survival probability comes from the response to the following question:

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

A similar question was asked in wave 2 except that respondents were asked to report the chances on a zero to 100 point scale.[8] We rescaled the responses so that they lie between zero and one, and we treat

---

[5] The survey over-sampled blacks, Hispanics and Floridians, which can be accounted for by weighting.

[6] Of those who were not interviewed, 225 are known to have died between the waves, 692 are known to be alive, and the mortality status of the remaining 243 is unknown. Many thanks to Janet Keller for supplying us with data on the mortality status of the HRS population.

[7] The survey is described more fully in Juster and Suzman (1995).

[8] We discuss the purpose of this change in the following section.

them as probabilities.

Table 1 shows the means of our variables by survivorship status in wave 2. Among our sample of 11,090 respondents, 183 died, 10,642 survived, and the status of 265 others was unknown. The 183 deaths out of 10,825 known outcomes implies a two-year mortality rate of 0.0169.[9]

The survivors gave considerably higher subjective survival probabilities than those who died. Therefore, at least in a gross way, the subjective survival probabilities predict mortality. Comparing the mean values of variables for those who died with those who survived, we see that the survivors are younger by about a year on average, and they had higher income and assets than those who died, reflecting differential mortality by economic status. As would be expected, more men died than women: men comprise about 60 percent of the deceased and only 45 percent of the survivors. This difference by sex is due in large part to the higher mortality rates faced by men, but also reflects differences by sex in the age distribution of the sample: the mean age of women in the HRS was 54.1 compared to a mean age of 57.4 for men, which is the result of men marrying younger women.

Singles died at a greater rate than married respondents. Those who died were in much worse health at baseline than the survivors. The fraction in poor health in each of the two groups was 0.36 and 0.06. Smoking is a risk factor, and there is a larger fraction of smokers among those who died than among those who survived. Disease conditions all have the expected relationship: among those who died greater fractions had each of the disease conditions, and in some cases the differences are large. For example, 23 percent of those who died had been told by a doctor at some point that they had a cancer, whereas only 5 percent of those who survived had been told of a cancer.[10]

The averages over those whose survivorship status is unknown show that they are similar to the survivors in some respects such as age, sex and the subjective survival probabilities; yet they are similar to the deceased along other characteristics such as income and marital status. Their health as measured by self-assessed health status and disease conditions at baseline is better than those who died, but less good than the health of those who lived. Therefore we anticipate that the mortality rate for this group will be somewhat higher than for the sample as a whole.[11]

---

[9]Later in this paper we will compare a mortality rate calculated over just the population-representative part of the sample with mortality rate calculated from a life table; as we will see, the two are rather close.

[10]The questions about disease condition are mostly of the form: "Has a doctor ever told you you had..." Thus the disease need not be current, and in some cases may have been completely cured.

[11]Later in this paper we will estimate a predicted mortality rate for this group based on observable characteristics and we find this to be the case.

Table 2 shows the average reported probability of survival to age 75 (which we term P75) and to age 85 (P85) in wave 2 among the population-representative portion of our sample (those born in the years 1931 to 1941). We compare this average with the survival probability calculated from a 1990 life table.[12] The average subjective survival probability is remarkably close to the life table average, which is what we found in wave 1. The reported survival to 85 is considerably higher than the life table rate, implying that respondents overestimate the conditional probability of survival to 85 given survival to 75. Hamermesh (1985) finds similar results with individuals slightly underestimating short-term survival probabilities and over-estimating longer-term probabilities relative to life table values. He views this over-estimate as possible evidence that individuals "extrapolate past increases in longevity" (p. 393).

Women give higher averages than men, as they should, although the difference is smaller than the life table difference. A possible explanation is that when forming expectations, individuals take into account the actual mortality experience of those around them, including both men and women, and fail to adjust fully for differences between the sexes. Men therefore overestimate, and women underestimate their survival probability.

3 Comparison of Subjective Survival Probabilities Across Waves

In wave 1 a rather large fraction of respondents gave focal-point responses to the subjective survival questions: of those in our sample, 5.8 percent reported 0 for P75, 21.2 percent reported 0.5 (actually 5 on the 0-10 point scale) and another 21.2 percent reported 1.0. Thinking that the zeros and ones were partly due to the limited scale, the survey designers rescaled the probability questions in wave 2 to range from 0 to 100.[13]

By several measures of internal consistency, the wave 2 question is superior to the wave 1 question (table 3). Of those who responded to the survivorship questions in both waves, a greater number in wave 2 gave P75 > P85 than in wave 1, 74.8 percent compared to 72.9 percent. If the answers are true probabilities then P75 will be greater than P85. There were also smaller fractions in wave 2 at the focal points of zero and 1.0 on both questions. However this improvement was at least partially offset by an

[12]The life table averages come from weighting the single-age survival rates by the weighted number of HRS respondents by age and sex.

[13]Someone with a subjective survival probability of 0.96 might choose 1.0 when faced with the choice of 0.9 or 1.0.

increase in the fraction giving the focal response of 0.5 in the second wave. Possibly people are used to thinking in terms of a "50-50 chance" rather than "5 out of 10," so that they respond to uncertainty with a "50" on a 0--100 point scale more frequently than with a "5" on a 0--10 point scale.

Table 4 shows the stability of the responses across waves at the individual level. Even though the scale changed, a rather large fraction, 26.9 percent, gave the same response in both waves, mostly at 0.5 or 1.0. More respondents gave a decline in the probability than gave an increase; yet, if survival expectations follow a life table, they must, by construction, increase with age. The decline in survival probabilities may therefore represent a change in awareness of risk factors or a general increase in pessimism. Of course, the change in the scale itself could have caused the decline.

Figure 1 shows the distributions of responses. In wave 2 there was some success in getting a more even distribution of responses than in wave 1: about 16 percent of respondents gave a wave 2 response that is not divisible by 10 on the 0--100 point scale. However, most of these are at 75: the new scale apparently created a new focal point by drawing responses from 7 and 8.

Figure 2 shows the distribution of wave 2 responses conditional on a wave 1 response of zero, 0.5 or 1.0. Among those who gave a zero in wave 1, about 36 percent stayed at zero, a much greater percentage than the unconditional rate (5.2 percent). About 23 percent gave 0.5 and about 7 percent gave 1.0. These kinds of individual transitions are reasonable in an uncertain environment where new information can cause an upward revision in an individual's optimism. However, some of the movement is undoubtedly due to observation error.

The second panel in the figure shows the distribution conditional on a wave 1 response of 0.5. Forty-seven percent of the individuals also report in wave 2 a survival probability of 0.5, while 10 percent increased their response from 0.5 to 1.0. The bottom panel repeats the analysis for individuals initially reporting 1.0. Forty-four percent of the population stayed at exactly 1.0, and most of the rest dispersed to 0.9, 0.8, 0.75 or 0.5.

The overall impression from figure 2 is that conditioning on wave 1 has strong effects; the conditional distributions are substantially different from the unconditional distributions and large numbers of individuals give the same response in the two waves. It also shows that expanding the scale in wave 2 did not induce many to move from focal responses to other nearby values.

In our previous work, we found that the levels of P75 and P85 are correlated with a number of observable characteristics such as disease conditions, lack of exercise, and smoking status which are associated with elevated mortality risk in the population. Table 5 has similar results for several characteristics using the wave 2 data. Mortality rates are known to decrease with income (Kitagawa and

Hauser, 1973; Caldwell and Diamond, 1979). Here we see that P75 increases monotonically with both income and wealth. Moving from the lowest wealth quartile to the highest, the subjective probability of surviving to age 75 increases from 0.57 to 0.70.[14] As shown in table 1, this variation by income or wealth is qualitatively a correct predictor of actual survivorship between waves 1 and 2.

As was the case with the mortality differential by income, the known negative relationship between schooling level and death rates (Rosen and Taubman, 1979; Zopf, 1992) is mirrored in differences in P75 and P85 with, for example, P75 ranging from 0.56 for those with less than a high school education to 0.69 for those with a college degree. The variation by smoking status also accords with known patterns (Preston, 1970), smokers report significantly lower survival probabilities than do non-smokers. Those who quit smoking report survival probabilities that are almost identical to those of non-smokers. The patterns observed for P85 are similar to those for P75.

4 Changes in Survival Probabilities

According to the results in table 4 and the conditional distributions in figure 2, there is considerable movement in the subjective survival probabilities at the individual level. The most obvious explanation for changes in P75 across waves is an unanticipated change in health. The top panel of table 6 shows that there is substantial change in self-assessed health status between waves. As shown along the diagonal, only about half of the respondents gave the same reported health status in wave 2 as in wave 1. In wave 1 the average response was between good and very good, so the table indicates regression towards the mean. For example, among the those in good health in wave 1, about 30 percent reported better health in wave 2 relative to wave 1, and 17 percent reported worse. These kinds of transitions could be the result of reporting error, or of transitory health status.

The bottom panel shows the change in P75 corresponding to each cell in the top panel. Among those whose health remained the same (the diagonal) the survival probability decreased slightly, and several of the changes are significant. As discussed in connection with table 4, we have no good explanation for the decline. Below the diagonal, health worsened between the waves, and in all cases P75 decreased as well. Some of the declines were very large, especially those associated with a decline to poor health. For example, among those in good health in wave 1, 2.5 percent reported poor health in wave 2,

[14]Based on wave 2 data, estimates of the linear regressions of the levels of P75 and P85 on a number of observed characteristics are in Appendix table A-1. These results are very similar to results based wave 1 data.

and their average subjective survival probability declined by 0.158. This is a large change, about three-fourths of the difference in P75 at baseline between those who survived and those who died (table 1). Entries above the diagonal correspond to an improvement in health, and with one exception (the transition from poor to very good) the changes in P75 were positive. For example, 2.1 percent of those in fair health reported excellent health in wave 2 and their survival probability increased by 0.188. We conclude that the relationship between health change in the panel and the change in the survival probability is qualitatively the same as the relationship between health and the survival probability in cross-section.

We now examine the factors associated with changes in the subjective probabilities. When individuals make projections about their survival probability, one would imagine that they incorporate expectations about future changes in health and other factors contributing to survival. Thus it ought to be unexpected changes or new information that most affects reported survival probabilities. This new information would be expected to also affect subjective health status. We therefore examine the change in P75 as a function of characteristics that are at least partly unexpected, including the death of a parent or the onset of disease.

Focusing on changes in P75 rather than levels also allows us to control for unobserved differences across individuals. In cross-section P75 varies in a reasonable way with a number of observable characteristics, such as the frequency of exercise, disease conditions, and smoking status (Hurd and McGarry, 1995). However, this kind of variation does not imply causality. It may be that there exist unobserved measures of underlying healthiness and optimism that are correlated with both reported life expectancy and the observable characteristics. In the panel we can specify a relationship that can reasonably be interpreted to be causal because we can relate changes in the subjective survival probability to changes in observable characteristics that are at least partly unexpected, for example the death of a parent or the onset of disease.

Table 7 has the coefficients and standard errors from the regression of the changes in P75 and P85 (wave2 - wave1) on changes in the survivorship of the respondent's parents and on onset of disease conditions.[15] The average change in P75 is -0.0149 and the change in P85 is -0.0217.

In a cross sectional regression controlling for many characteristics, respondents whose parents were alive or whose parents died at advanced age reported higher subjective survival probabilities than respondents whose parent died at younger ages. As might be implied by these relationships, changes in

[15]The regression can be interpreted as one that controls for individual effects: differencing P75 and P85 removes an individual effect.

the mortality status of parents has a numerically large effect on changes in survival probabilities, although few coefficients are statistically significant because few parents died.[16] If the respondent's mother died between waves and was younger than 75 at her death, the respondent reduced the probability of survival to 75 by 0.12.[17] This change is rather large given an average P75 of about 0.64. If the mother died at 75 or older the survival probability was not reduced significantly. The effects a father's death are similar: if the father died before age 75, P75 is reduced by -0.124, but there was little effect if he died at 75 or over.

We found in cross-section that P75 and P85 were affected differentially by the age of the parents' death: an early death (before 75) affected both P75 and P85, a death between 75 and 84 affected P85 but not P75, and a death after 85 affected neither. Comparing the effects of the mother's death between 75 and 85 on P75 with the effects on P85 points to some suggestions of a differential here. Overall, however, the effects are considerably weaker than in cross-section, and there is no such pattern for the effects of the father's death.

One might imagine that the death or a mother or the death of a father have differing effects depending on the sex of a child, and, indeed, in wave 1 cross-section we did find such a difference. We allow for differing responses by sex in both equations by including the interaction of a dummy variable indicating the death of a mother or father, with the dummy variable indicating that the respondent is male. Thus if the respondent were male and his mother died aged 85 or over the effect on P75 would be -0.031 + 0.036, which is practically zero. However, because the coefficients on the two interaction terms (mother died and respondent male, father died and respondent male) are almost the same, the interactions reduce to a single additive effect implying that males reduce P75 by less when a parent dies than do females. On P85, however, the effect is strong, and it is consistent with the cross section results. If the respondent is female the death of a mother has large and significantly negative effects on P85, while the effects of a mother's death of male respondents are significantly smaller (-0.197+0.096, or smaller depending on the age of the mother).

The effect of a parent's death may operate through both biological and psychological mechanisms. If the parent died of a cancer that is known to have a genetic link, the child might correctly reassess his own life expectancy. However, a parent's death may also affect the respondent's reported probability because it reminds him of his own mortality. Because there are no genetic links between spouses, one

---

[16]357 (or 4.2 percent) of mothers and 255 (or 3.0 percent) of fathers died between the waves.

[17]The reference group is respondents whose parents survived between waves.

would expect the impact of the death of a spouse to be largely psychological (although environmental factors may impact both spouses). We find that the death of a spouse has a large and significantly negative effect on the survival probability to age 75, and a smaller, insignificant effect on survival to 85.[18] In contrast to the effects of parental and spousal deaths, the death of a sibling has no effect in either equation, although siblings share similar genetic makeup.[19] Perhaps respondents are less close psychologically to siblings than to parents or a spouse, and therefore less affected by the death of a sibling.

The remainder of table 7 has effects associated with the onset of disease between the two waves. All the coefficients in each equation are negative, although, with the exception of the effects of cancer, they are not significantly different from zero, most likely because of the small number of observations involved.[20] Nonetheless, the results show that respondents appropriately reduced their subjective survival probabilities at a new diagnosis, particularly for conditions that are more life threatening such as cancer, heart conditions, and lung disease.

The subjective probability of survival is highly correlated with self-assessed health status, but, in principle, it includes an additional expectational component. For example, if a respondent is healthy today but some event occurs that increases the likelihood of disease sometime in the future, that event should reduce P75 and P85 but not self-assessed health. The death of a parent may be such an event. Except for the stress of bereavement, it is difficult to think that the death itself could affect the respondent's current health status, but it may increase the subjective likelihood of onset of a genetically-linked disease.[21] We test this idea in our data by finding whether the change in subjective health status across waves is related to a parent's death or the death of a spouse. We use a multinomial logistic model which is more appropriate for self-assessed health than a linear regression because the health measure is categorical. We defined three health states in wave 2 relative to wave 1; improved, stayed the same, or declined. Approximately 20 percent of the sample had an improvement in health, 53 percent had no change, and 27 percent had a worsening of health (table 8). A positive coefficient in the multinomial logit regression

---

[18]Most of the spouses who die are male: 82 husbands died and 18 wives died. An interaction of the death of a spouse with the sex of the respondent was not significant and is excluded from the final regression.

[19]111 (or 1.3 percent) of respondents had a sibling die between waves.

[20]The number of new cases is small: for example, about 115 or 1.4 percent of the sample was newly diagnosed with cancer, and 233 or 2.7 percent with heart conditions. If all conditions other than cancer are combined into one measure of "other disease conditions" they affect P75 significantly at the 10 percent level, and P85 at the 1 percent level. The effect of cancer on the variables is unchanged.

[21]We would like to use data on the cause of the parent's death, but those data are not in the HRS.

means that the variable increases the probability of the corresponding health change. If a parent's death resulted in a worsening of the respondent's self-assessed health, the coefficients under "health better" should be negative and the coefficients under "health worse" should be positive. We find no evidence of this effect. For example, if the mother died between the waves and her age at death was less than 75, the respondent was more likely to have an improvement in health between waves (coefficient of 0.367) *and* more likely to have a worsening in health (coefficient of 0.520), than to have stable health status, although neither of the coefficient estimates is significantly different from zero. If the mother died between the ages of 75 and 85, the probability of an improvement in health does fall and the probability of worsening of health increases but the effects are again insignificant. All coefficients for the death of a father act to reduce the probability of changing states, and none is significantly different from zero. The death of a spouse increases the probability of an improvement in health status and decreases the probability of a worsening of status, but again, insignificantly.[22]

In contrast to these weak and contradictory effects, the onset of a disease affects current health status in the expected direction. The coefficients on the disease measures typically increase (significantly) the probability of moving to worse health, and in most cases, decrease the probability of improved health, or else have no significant effect.

Because diseases lower both reported health status and the reported survival probability, while the death of a parent or of a spouse significantly lowers only the reported value of P75 and P85, we conclude that the subjective survival probabilities measure more than health status: they have an expectational component as well as a health-status component.

5 Mortality Outcomes

Before we study the relationship between mortality and the subjective survival probabilities, we compare the mortality experience of the HRS respondents to the mortality experience of the population as found in life tables. To make a meaningful population comparison, we restrict our sample to respondents born from 1931 through 1941. Table 9 shows the mortality experience of the wave 1 sample and the number

[22] We also estimated a linear regression for change in health status where the change was measured as the difference between wave 2 status and wave 1 status and health is measured on a scale of 1-5. Taken as a group the variables indicating the parents' vital status were not significant.

whose mortality status is unknown.[23] The counts are weighted to control for the over-sampling of blacks, Hispanics, and Floridians.

As estimated from the numbers in table 9, the mortality rate for the sample is 0.0149 which is somewhat lower than the mortality rate calculated from the life table, 0.0181. For women and for men the rates from the HRS are 0.0111 and 0.0194 compared with 0.0134 and 0.0236 from population life tables. There are several explanations for the consistently lower mortality among respondents. First, the HRS is representative of the *non-institutional* population. Those in institutions most likely have higher mortality risk than the non-institutionalized population. Thus mortality in the HRS will be lower than the national average. Second, as suggested by the covariates in table 1, the mortality rate of the 164.5 (weighted) cases whose mortality status is not known is likely to be above average. Their inclusion would increase the average HRS mortality rate. Third, the time span between waves 1 and 2 is not exactly 2 years. The mean interval is 22.5 months, and the modal interval is 22 months. Increasing the length of the interval to 24 months (about 7 percent) would increase the number of deaths in the HRS. Finally, the difference between the two rates can be further reduced if a 1993 life table is used in lieu of a 1990 life table, thus incorporating recent improvement in longevity. We will make several of these adjustments in results to be reported later in this paper.

We now examine how well the subjective survival probability at wave 1 predicts actual mortality in the panel for our sample of individuals age 46-65 whose mortality status is known. We have already seen that P75 predicts mortality on average because in table 1 those who died had substantially lower subjective survival probabilities. Figure 3 shows the cumulative distributions of P75 for those who died compared with those who survived. Not only is the average different for these two groups, but the differences persist throughout the distributions. For example, about 11 percent of those who survived reported P75 to be 0.40 or less whereas 43 percent of those who died gave a value of 0.40 or less. The median values for the two distributions were 0.7 and 0.5.

Figure 4 shows two-year mortality rates as a function of P75 and P85. Mortality rates decline almost monotonically as P75 varies from 0.1 to 1.0. Although there was considerable bunching of responses at 0.5 (Figure 1) the mortality rate at 0.5 is not noticeably different from mortality rates in the range of 0.3 to 0.7, suggesting that respondents who report a value of 0.5 are drawn from nearby

[23]We have excluded observations on respondents who were interviewed by proxy because we want to use P75 as an explanatory variable for mortality; yet proxies were not asked about the respondent's subjective survival probability. Including proxy responses increases the HRS two-year mortality rate for women from 0.0111 to 0.0118 and the rate for the entire sample from 0.0149 to 0.0154. The mortality rate for men is unchanged.

probability points. Mortality at 0.1 is very high, implying that those respondents understood both the probability question and its implications. Mortality at zero is greater than at any other point except 0.1. We interpret this to mean that zero contains a mixture of individuals who appropriately answered the question with those who did not. Therefore the mortality outcome at zero is a mixture of a high rate for one group and, a rate perhaps closer to average, for another group.

The risk curve for P85 is considerably shallower, and even has an increase from 0.7 to 1.0. An implication is that P85 contains more observation error than P75. This implication is consistent with the results in table 2 where the average of P85 was considerable higher than the life table average. An explanation is that survival to 85 is a very distant event from the point of view of people in their 50s.

In table 10 we examine whether P75 has any power in predicting mortality after controlling for other observable characteristics. We estimate the partial effect of P75 from a probit model where the outcome is mortality between waves and the explanatory variables include risk factors such as socio-economic status, health behaviors and disease conditions. The average mortality rate over this group was 0.0167, which is somewhat higher than the rate over the age-eligible sample (in table 9), due to the addition of those between the ages of 62 and 65. The table reports the probability derivatives of the coefficient estimates, and the statistic for testing the null hypothesis that the effect is zero.[24] The table has two sets of results, those with and without subjective health status. We will first discuss those that exclude health status.

The effect of the subjective survival probability is significantly different from zero and fairly large. An increase in the probability from zero to 1.0 reduces the mortality hazard by 0.013, which is about three-quarters of the average hazard. In a mortality model such as a proportional hazards model this change would reduce considerably the likelihood of survivorship to advanced age. As will be seen later in the table the effect is about as large as that associated with moving from the lowest income quartile to the highest and almost as large as the effect associated with smoking.

Mortality falls with income, and particularly in the lowest income quartile the effect is strong, increasing the mortality rate by 0.010 or about 60 percent. Mortality falls with wealth through the third quartile, but increases at the fourth (highest) quartile. We have no explanation for this nonlinearity. The mortality rate increases with age, and the difference between the risk of a 51 year-old and a 61 year-old is about 0.01. Marital status per se has no effect. Apparently the difference by marital status observed

---

[24]The square-root chi-square statistic is asymptotically standard normal under the null hypothesis. We take a value of 1.96 to indicate significance in a two-sided test.

in table 1 is the result of differences in other variables that are correlated with marital status. Men have mortality rates about 0.011 higher than women, which is about the same as would be found in a life table for people of the HRS age range. Thus, controlling for all the covariates in table 10 does not reduce the male-female mortality differential.

Physical activity has large effects. For example, never having light physical activity increases the morality rate by 0.015, the same amount as does smoking. However, the significant distinction is between those who never have light physical activity and those who have any at all. Examples of light physical activity are walking, dancing, gardening, golfing and bowling, so the question may simply classify individuals as to those who cannot have any physical activity (and are in very poor health) and those who can. The effect of heavy physical activity only appears in the comparison of those with three or more sessions per week, suggesting that many completely healthy people never engage in any heavy physical activity.

Whites have slightly lower mortality rates than non-whites after controlling for other risk factors but the effect is not significantly different from zero. The difference by race that is evident in the raw data (table 1) is adequately explained by the observable factors that are included in the probit specification. Smoking approximately doubles the mortality rate, even after controlling for many other risk factors. Former smokers also have an elevated risk, with a mortality rate that is about 65 percent higher than non-smokers, although note that in table 5 they report no difference in expectations. Moderate drinking is thought to be associated with increased longevity in epidemiological data (Ellison, 1990), but in the regression the difference is zero.

Among the disease conditions, cancer is the strongest predictor of mortality, approximately doubling the two-year mortality rate.[25] The reports on the subjective survival probabilities are consistent with this result: in table 7, a new cancer had the strongest effect among the disease conditions on reducing the subjective survival probability. In a similar way, having had a heart attack, heart failure and having had a stroke all approximately double mortality risk. New diagnoses of heart attack and heart failure are strong determinants of a decline in the subjective survival probability in table 7.

The right two columns show comparable results when subjective health status is included. Fair or poor health predicts large increases in mortality rates with each approximately doubling the rate compared with the preceeding, better health status. Including health status reduces the effect of P75 by

[25]The indicator for cancer comes from the baseline interview, and indicates that a doctor at sometime told the respondent of a cancer. Thus, the cancer does not have to be active.

38 percent so that it is no longer significant, but it has practically no effect on the other coefficients. All of the other included variables with the exception of the subjective probability are, in principle, objective. These results show that individuals can provide information about health status that is approximately orthogonal to objective information.

Even though self-assessed health status reduces the effects of the subjective survival probability, we do not believe that self-assessed health status should replace subjective survival probabilities in models of economic behavior. First, the models call for the probabilities, not health conditions that may be predictive of mortality, and they call for the beliefs of the decision maker about the probabilities. Second, as we have seen, the subjective probabilities have an expectational component that is absent from health status. The models are based on the present situation and on anticipated changes.

As shown in table 9, our wave 1 age-eligible sample has 164.5 weighted cases whose survivorship status in wave 2 is not known. We can use estimates from the probit model of mortality to predict average mortality among these respondents, and adjust the overall mortality rate of the wave 1 sample accordingly. Using the average values from wave 1 of the explanatory variables for the missing cases we predict an average two-year mortality rate to be 0.0247, which is 66 percent higher than the mortality rate of the wave 1 sample whose survivorship is known. Based on this estimate we can adjust upward the total mortality rate of the wave 1 sample from 0.0149 as in table 9 to 0.0152. We can also increase the HRS mortality rate to account for the 22.5 month average interval between interviews (rather than 24 months) by 6.7 percent to 0.0162.

Our best comparison with life table mortality rates comes from interpolating between a 1990 and a 2000 life table to construct a 1993 life table. The predicted mortality rate from our 1993 life table is 0.0174 for the age-eligible sample. Therefore the final comparison with the life table shows a difference of 0.0012 with a standard error of 0.0015 (coming entirely from the standard error of the HRS sample). Thus, the difference between the two rates is not significantly different from zero.[26]

6 Conclusion

Previous research had established that respondents can and will answer questions about subjective

[26]We have no good method for adjusting the HRS mortality rate for the exclusion of the institutionalized population. Were we able to make such an adjustment the difference between the life table and the HRS mortality rates would be reduced further.

probabilities;[27] that the survival probabilities aggregate to averages in life tables, and that they covary with risk factors in an appropriate way. The objective of this paper was to take the essential step of understanding how subjective probabilities evolve in response to new information and how well they predict mortality. In doing so we hope to have increased confidence in these measures so that they may eventually used to understand and predict economic behavior.

We found that subjective survival declines with the death of a parent, but that self-assessed health is not affected. We interpret this to mean that the subjective survival probability has an expectational element and that it is not simply an alternative measure of health status. In addition, the survival probability predicts mortality. Those who survived from wave 1 to wave 2 of HRS gave subjective survival probabilities in wave 1 that were about 50 percent higher than those who died between the waves. Also, while it is apparent that the probabilities contain some observation error (as evidenced by the bunching at 0.0, 0.5 and 1.0), we have no reason to believe that they contain any more observation error than other types of economic data such as assets. Nonetheless, models that use subjective probability data to estimate behavioral relationships need to account for observation error.

Several steps remain before we can use subjective probabilities of survival in models. First, we need to find if they are correctly scaled numerically, and if not, how to adjust the scaling. Because the probability refers to survival to age 75, while observed mortality is over two years, assessing the scaling requires a mortality model, such as a proportional hazards model, and the associated distributional assumptions. Second, we need to construct survival curves because any behavioral model will require survival probabilities to all ages, not just the target age of 75 or 85. This step will also require a mortality model.

Once these steps have been accomplished, the survival probabilities can be used as explanatory variables in models. For example, in a life-cycle model individuals with greater survival probabilities should work longer and save more. The subjective survival probabilities will provide individual variation in survival that cannot be obtained from life tables. This variation will sharpen parameter estimates, and help to explain heterogeneity in behavior.

---

[27]The response rate in wave 1 of HRS on subjective survival was 98 percent, which is considerably higher than on questions about intentions such as the age of retirement.

References

Caldwell, Steven and Theodore Diamond, 1979. "Income Differentials in Mortality: Preliminary Results Based on IRS-SSA Linked Data," in Linda DelBene and Fritz Scheuren, eds., *Statistical Uses of Administrative Records with Emphasis on Mortality and Disability Research* (Washington DC: U.S. Department of Health Education and Welfare).

Dominitz, Jeffery, 1997. "The Distribution of Income Uncertainty Across the U.S. Labor Force: Estimates from Expectations and Realizations Data." mimeo: California Institute of Technology.

Dominitz, Jeffery and Charles Manski, 1994. "Using Expectations Data to Study Subjective Income Expectations." NBER working paper 4937, Cambridge, MA (1994).

Ellison, R. Curtis, 1990. "Cheers," *Epidemiology* 1: 337-339.

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): 911-914.

Hurd, Michael D., Daniel McFadden and Li Gan, 1995. "Subjective Survival Curves and Life Cycle Behavior," presented at the NBER Conference on the Economics of Aging, The Boulders, May, 1995, revised June, 1996.

Hurd, Michael D., and Kathleen McGarry. 1995. "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study." *Journal of Human Resources* 30: s268-s292.

Juster, F. Thomas. 1966. "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design." *Journal of the American Statistical Association* 61: 658-696.

Juster, F. Thomas and Richard Suzman. 1995. "An Overview of the Health and Retirement Study."

Journal of Human Resources 40: s7-s56.

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

Manski, Charles F. 1990. "The Use of Intentions Data to Predict Behavior: A Best Case Analysis," *Journal of the American Statistical Association*, 85: 934-940.

Preston, Samuel H., 1970. "The Age-Incidence of Death from Smoking," *Journal of the American Statistical Association* 65(331): 1125-1130.

Taubman, Paul and Sherwin Rosen, 1979. "Changes in the Impact of Education and Income on Mortality in the U.S.," in Linda DelBene and Fritz Scheuren, eds., *Statistical Uses of Administrative Records with Emphasis on Mortality and Disability Research* (Washington DC: U.S. Department of Health Education and Welfare).

Zopf, Paul E. Jr., 1992. *Mortality Patterns and Trends in the United States*, Westport, CT: Greenwood Press.

Figure 1. Distributions of P75

Waves 1 and 2

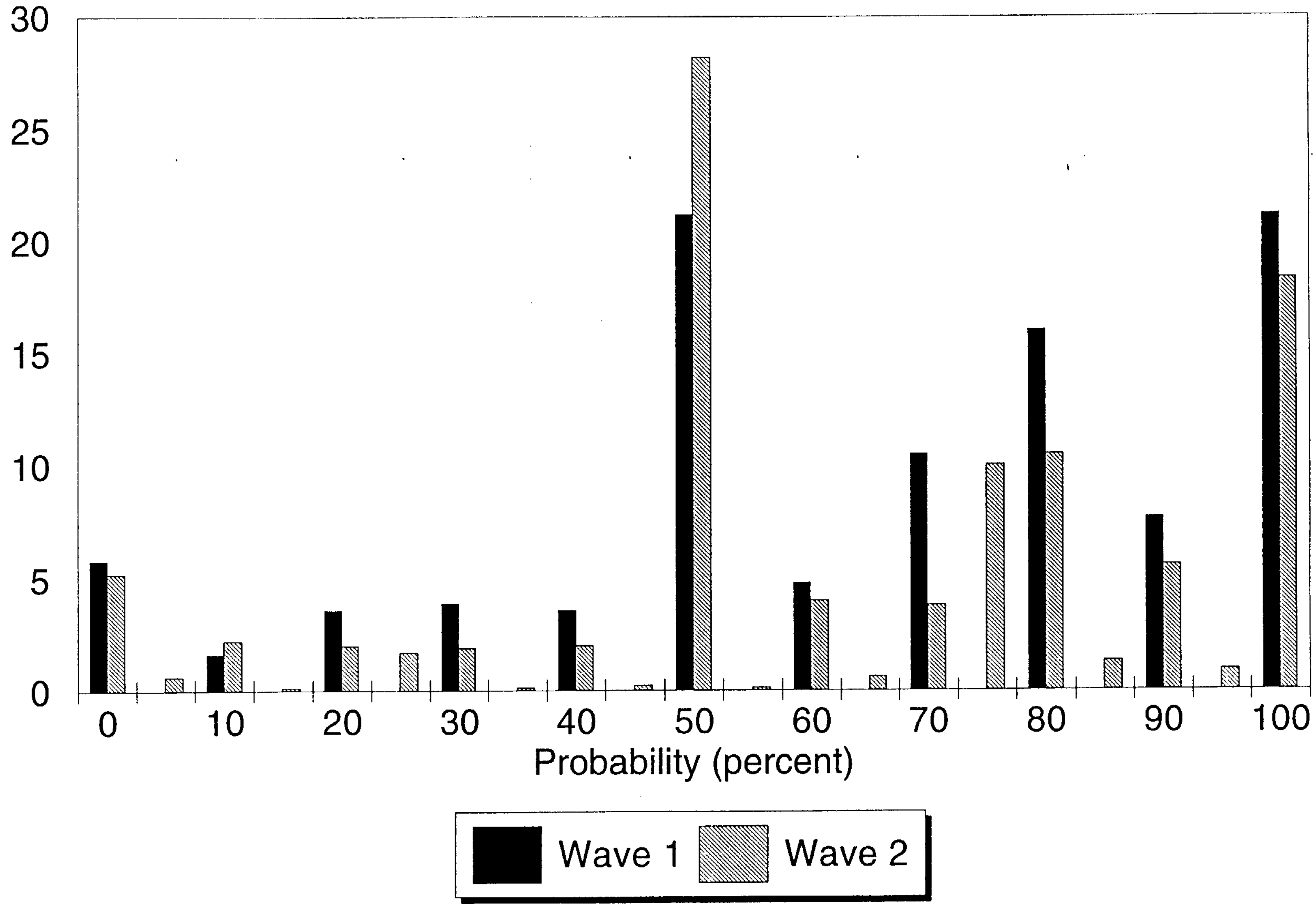

Figure 2. Wave 2 subjective probability of survival to 75 conditional on wave 1 response

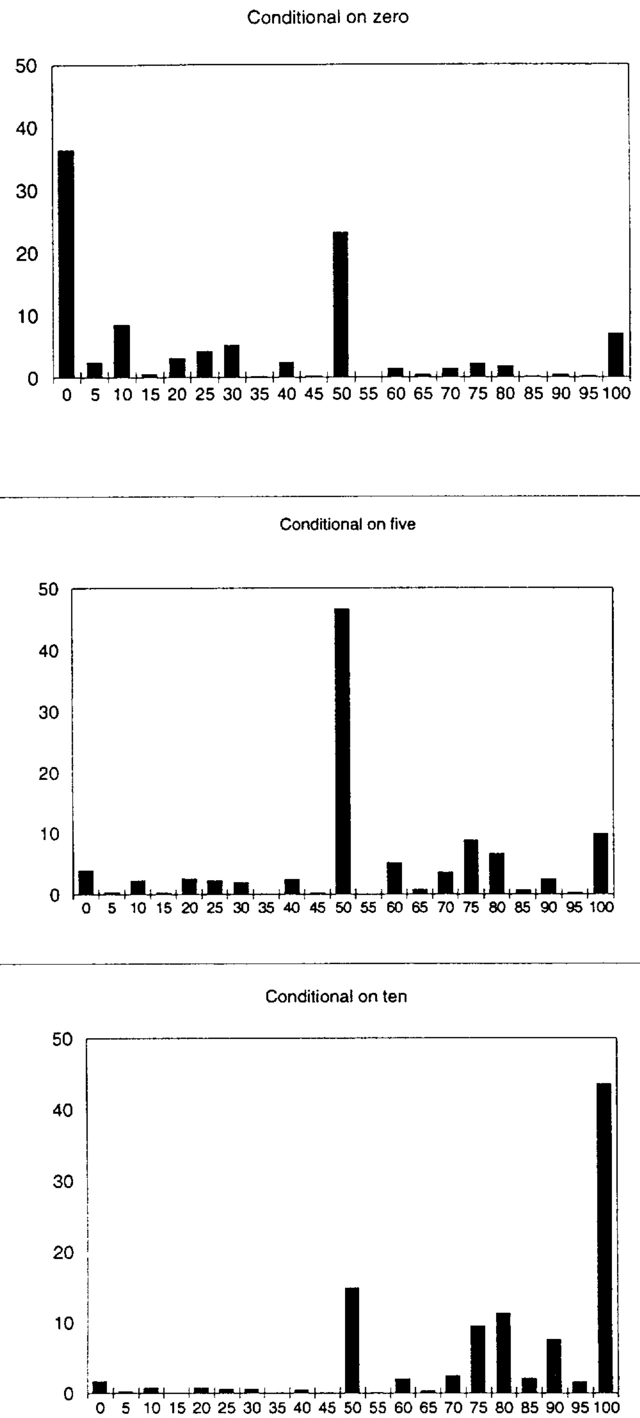

Figure 3. Subjective survival

Cumulative distributions

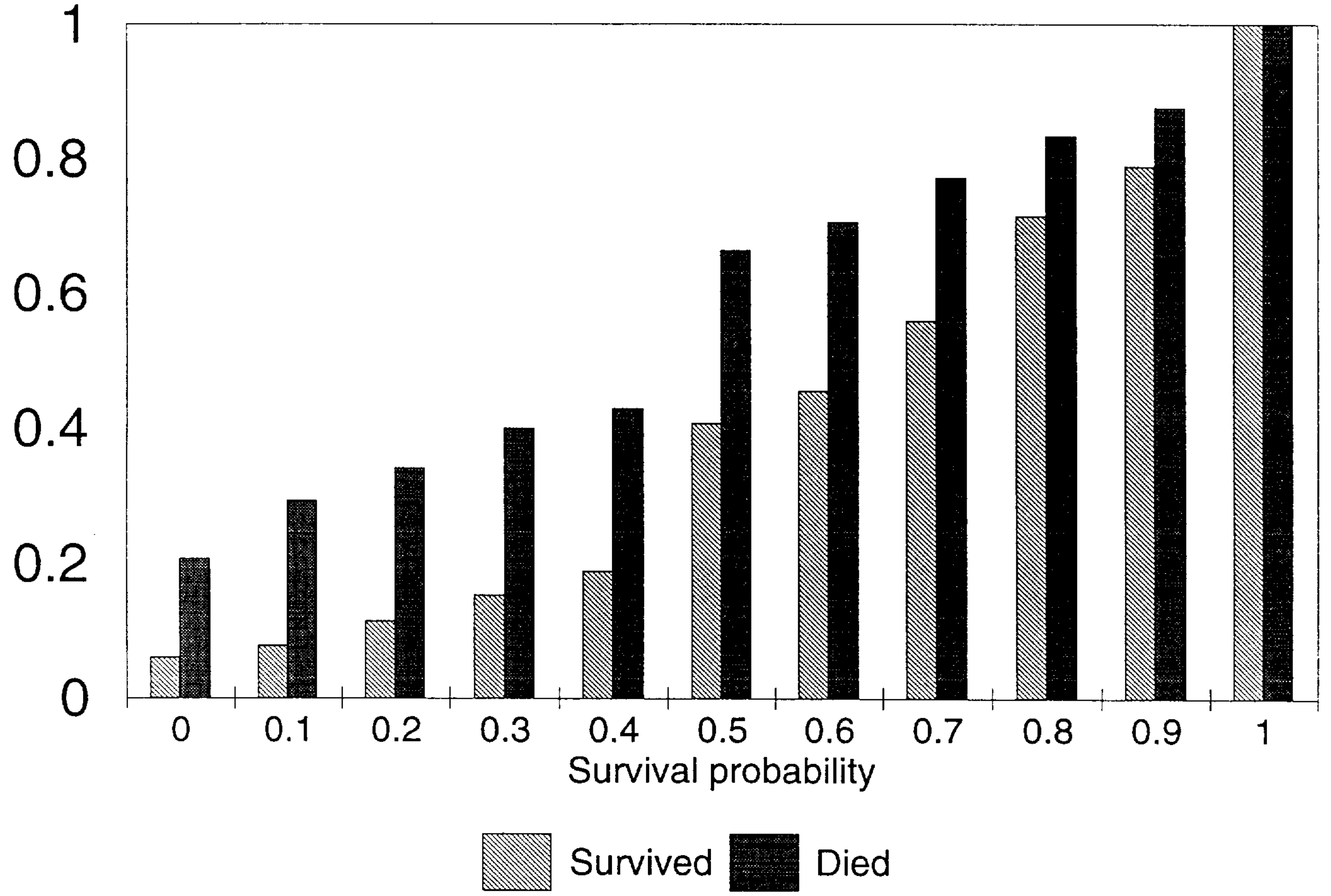

Figure 4. Two-year mortality rate

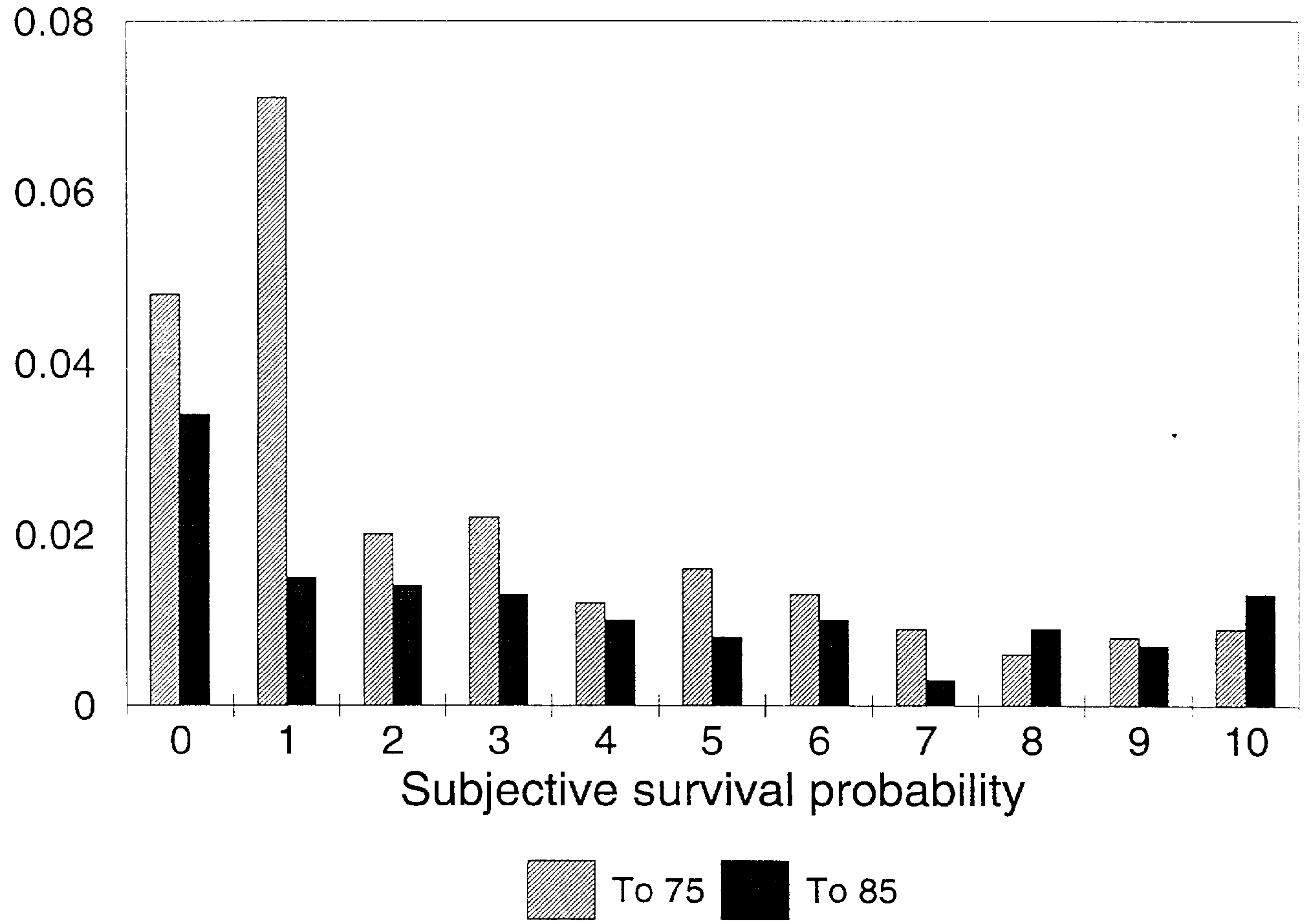

Table 1
Means of wave 1 variables by survivorship to wave 2

	Died between waves		Survived to wave 2		Survivorship unknown	
Variable	Mean	Std Err	Mean	Std Err	Mean	Std Err
Prob live to 75	0.45	0.03	0.65	0.00	0.66	0.02
Prob live to 85	0.28	0.02	0.43	0.00	0.42	0.02
Age	56.91	0.30	55.71	0.04	55.95	0.23
Family income (thousands)	0.32	0.07	0.46	0.01	0.28	0.02
Wealth (thousands)	0.17	0.03	0.28	0.01	0.22	0.03
Male	0.60	0.04	0.45	0.00	0.44	0.03
Married	0.67	0.03	0.79	0.00	0.61	0.03
Self reported health						
excellent	0.08	0.02	0.24	0.00	0.20	0.02
very good	0.13	0.02	0.30	0.00	0.23	0.03
good	0.18	0.03	0.28	0.00	0.28	0.03
fair	0.25	0.03	0.12	0.00	0.16	0.02
poor	0.36	0.04	0.06	0.00	0.13	0.02
Smokes	0.38	0.04	0.26	0.00	0.33	0.03
Nonwhite	0.22	0.03	0.14	0.00	0.29	0.03
High blood pressure	0.52	0.04	0.37	0.00	0.42	0.03
Diabetes	0.25	0.03	0.10	0.00	0.12	0.02
Cancer	0.23	0.03	0.05	0.00	0.05	0.01
Lung disease	0.18	0.03	0.08	0.00	0.09	0.02
Heart condition	0.35	0.04	0.13	0.00	0.14	0.02
Angina	0.11	0.02	0.04	0.00	0.06	0.01
Congestive heart failure	0.11	0.02	0.02	0.00	0.01	0.01
Stroke	0.10	0.02	0.02	0.00	0.03	0.01
Arthritis	0.43	0.04	0.38	0.00	0.41	0.03
Elementary	0.38	0.04	0.25	0.00	0.35	0.03
High School	0.39	0.04	0.37	0.00	0.27	0.03
College	0.23	0.03	0.38	0.00	0.38	0.03
Number of observations	183		10642		265	

Sample is individuals 46 to 65 in wave 1.

Table 2
Average probabilities of surviving to 75 or 85

	All		Women		Men	
	Age 75	Age 85	Age 75	Age 85	Age 75	Age 85
HRS wave 2 subjective probability*	0.645 (0.003)	0.427 (0.003)	0.663 (0.004)	0.460 (0.004)	0.622 (0.005)	0.388 (0.005)
1990 life table	0.690	0.356	0.756	0.444	0.608	0.247

* Weighted average of responses of individuals born from 1931 to 1941.

Table 3
Distributions of subjective survival probabilities

Probability comparison	Percent of HRS Respondents	
	Wave 1	Wave 2
P75 > P85	72.9	74.8
Both probabilities = 0	5.8	5.3
Both probabilities = 0.5	4.0	5.8
Both probabilities = 1.0	8.2	6.7
Both probabilities = some other value	7.0	5.7
P75 < P85	2.2	1.7
All	100.0	100.0

Sample is individuals age 46-65 in wave 1 who answered the probability question in both interviews.

Table 4
Distributions of subjective survival probabilities:
Comparison of wave 1 and wave 2 responses

Probability comparison	Percent of HRS Respondents	
	P75	P85
Both probabilities = 0	2.1	8.0
Both probabilities = 0.5	9.9	4.7
Both probabilities = 1.0	9.2	2.9
Both probabilities = some other value	5.7	6.1
Wave 1 probability > Wave 2 probability	39.1	42.8
Wave 1 probability < Wave 2 probability	34.1	35.4
All	100.0	100.0

Sample is individuals age 46-65 in wave 1 who answered the probability question in both interviews.

Table 5
Average subjective probability of surviving to age 75 (wave 2)

	Prob live to 75		Prob live to 85	
Characteristic	Prob	Std Err	Prob	Std Err
Income quartile				
lowest	0.58	0.007	0.391	0.007
second	0.62	0.006	0.389	0.006
third	0.65	0.006	0.418	0.006
highest	0.68	0.005	0.444	0.006
Wealth quartile				
lowest	0.57	0.007	0.377	0.007
second	0.62	0.006	0.394	0.006
third	0.65	0.006	0.415	0.006
highest	0.70	0.005	0.452	0.006
Schooling				
less than high school	0.56	0.007	0.367	0.007
high school graduate	0.63	0.005	0.397	0.005
college graduate	0.69	0.004	0.454	0.005
Smoking behavior				
never smoked	0.66	0.005	0.437	0.005
smoked but quit	0.65	0.005	0.416	0.005
current smoker	0.59	0.007	0.368	0.007

Sample is individuals age 46-65 in wave 1.

Table 6
Health transition probabilities and changes in subjective survival to age 75

	Health in wave 1				
Health in wave 2	Excellent	Very good	Good	Fair	Poor
	Transition probabilities				
Excellent	0.541	0.158	0.054	0.021	0.012
Very good	0.333	0.528	0.245	0.061	0.014
Good	0.105	0.263	0.526	0.278	0.083
Fair	0.017	0.043	0.149	0.502	0.327
Poor	0.005	0.008	0.025	0.138	0.565
All	1.000	1.000	1.000	1.000	1.000
	Change in survival probabilities				
Excellent	-0.017*	0.007	-0.002	0.188*	0.005
Very good	-0.026*	-0.021*	-0.000	0.062	-0.077
Good	-0.027	-0.036*	0.003	0.020	0.142*
Fair	-0.029	-0.094*	-0.038*	-0.007	0.047
Poor	-0.241	-0.220*	-0.158*	-0.062*	-0.016
All	-0.022*	-0.025*	-0.008	0.002*	0.018
Number of observations	1983	2452	2276	1089	526

Change in survival probabilities is wave 2 value minus wave 1. Overall average change is -0.014.
* denote significance at the 5% level. Sample is those age 46-65 in wave 1.

Table 7
Determinants of the change in the subjective probability of surviving (wave2-wave1)

	Living to 75		Living to 85	
Variable	Coeff	Std Error	Coeff	Std error
Intercept	-0.008*	-0.004	-0.014	0.004
Mother died between waves				
age at death < 75	-0.120*	0.043	-0.197*	0.046
75<= age at death <85	-0.014	0.025	-0.076*	0.027
85<= age at death	-0.031	0.029	-0.072*	0.031
Father died between waves				
age at death < 75	-0.124*	0.062	-0.094	0.067
75<= age at death <85	0.006	0.030	-0.002	0.032
85<= age at death	0.027	0.032	0.034	0.034
Respondent male and				
Mother died	0.036	0.031	0.096*	0.033
Father died	0.029	0.036	-0.038	0.039
Other mortality				
Spouse died between waves	-0.086*	0.029	-0.021	0.030
Sibling died between waves	-0.009	0.027	0.002	0.029
Onset of disease since wave 1				
high blood pressure	-0.012	0.015	-0.003	0.016
diabetes	0.000	0.021	-0.022	0.022
cancer	-0.112*	0.027	-0.099*	0.028
lung disease	-0.022	0.022	-0.022	0.023
heart attack	-0.026	0.020	-0.025	0.021
angina	-0.018	0.023	-0.034	0.024
congestive heart failure	-0.035	0.031	-0.021	0.032
stroke	-0.023	0.039	-0.045	0.040
arthritis	-0.008	0.012	-0.017	0.012
Number of observations	8509		8622	
Mean of dependent variable	-0.014		-0.022	

* Denotes significance at the 5% level.
Sample consists of individuals age 46-65 in wave 1.

Table 8
Multinomial logit coefficients: Health better (20%) or worse (27%) versus same (53%)

	Health better		Health Worse	
Variable	Coefficient	Std Err	Coefficient	Std Err
Intercept	-0.924*	0.031	-0.778*	0.029
Mother died between waves				
age at death < 75	0.367	0.396	0.520	0.336
75<= age at death <85	-0.367	0.249	0.061	0.193
85<= age at death	0.161	0.259	0.051	0.231
Father died between waves				
age at death < 75	-1.036	0.760	-0.527	0.531
75<= age at death <85	-0.228	0.293	-0.082	0.244
85<= age at death	-0.395	0.334	0.175	0.250
Respondent male and				
Mother died	-0.537	0.311	-0.114	0.242
Father died	-0.227	0.380	0.044	0.288
Other deaths				
Spouse died	0.234	0.234	-0.289	0.252
Sibling died	0.000	0.247	-0.036	0.224
Since wave 1 diagnosed with				
high blood pressure	0.146	0.143	0.619*	0.115
diabetes	0.197	0.187	0.341*	0.162
cancer	-0.073	0.313	1.272*	0.202
lung disease	-0.495*	0.234	0.192	0.163
heart attack	-0.036	0.208	0.936*	0.146
angina	-0.110	0.205	-0.391*	0.179
congestive heart failure	-0.567	0.325	-0.228	0.233
stroke	-0.218	0.405	0.728*	0.270
arthritis	-0.050	0.109	0.364*	0.089

Note: A positive coefficient increases the probability of a health change. Number of observations 8545. Sample is individuals 46-65 in wave 1.

Table 9
Weighted number of observations and two-year mortality rate, age 51-61 in wave 1

	All	Females	Males
Survived to wave 2	6887.75	3753.5	3134.25
Died before wave 2	104.0	42.0	62.0
Survivorship unknown	164.5	93.5	71.0
Mortality rate			
HRS estimate*	0.0149 (0.0015)	0.0111 (0.0017)	0.0194 (0.0025)
Life table	0.0181	0.0134	0.0236

* Life table estimate is weighted average (HRS weights) of single-age mortality rates from 1990 lifetable. Sample consists of those born in the years 1931-1941 and weighted to account for oversampling of blacks, Hispanics and Floridians.

Table 10
Determinants of mortality, wave 1 to wave 2
(average mortality rate = 0.0167)

	Excluding health		Including health	
	Effect on probability	Square root chi-square	Effect on probability	Square root chi-square
Subjective survival to age 75	-0.013	2.91	-0.008	1.69
Subjective health status:				
Excellent	--	--	0.001	0.16
Very good	--	--	0.003	0.49
Good	--	--	--	--
Fair	--	--	0.012	1.91
Poor	--	--	0.022	3.27
Household Income:				
lowest quartile (omitted)	0.010	2.00	0.008	1.58
2nd quartile	0.007	1.31	0.006	1.20
3rd quartile	-0.001	0.17	-0.001	0.18
highest quartile	--	--	--	--
Household Wealth:				
lowest quartile (omitted)	-0.001	0.21	-0.003	0.60
2nd quartile	-0.007	1.36	-0.007	1.49
3rd quartile	-0.010	2.02	-0.011	2.11
highest quartile	--	--	--	--
Age	0.001	2.07	0.001	2.15
Married	-0.003	0.73	-0.003	0.81
Nonwhite	0.005	1.31	0.004	1.11
Male	0.011	3.31	0.010	3.05
Frequency of light physical activity:				
3+ per week (omitted)	--	--	--	--
1-2 per week	-0.001	0.23	-0.001	0.16
1-3 per month	-0.008	1.23	-0.007	1.17
< 1 per month	0.009	1.72	0.008	1.52
never	0.015	3.78	0.013	3.19
Frequency of heavy physical activity:				
3+ per week (omitted)	--	--	--	--

1-2 per week	0.018	1.95	0.018	1.94
1-3 per month	0.023	2.52	0.024	2.53
< 1 per month	0.019	2.28	0.019	2.29
never	0.025	3.17	0.023	2.99
Smoking status:				
never smoked (omitted)	--	--	--	--
former smoker	0.011	2.78	0.011	2.77
smoker	0.015	3.77	0.015	3.74
Drinking behavior:				
does not drink (omitted)	--	--	--	--
drinks < once/day	-0.003	0.76	-0.002	0.49
1-2 /day	-0.001	0.22	0.001	0.11
3-4 /day	0.002	0.27	0.003	0.45
drink 5+	0.000	0.04	-0.002	0.17
Schooling:				
less than high school	-0.004	1.12	-0.005	1.45
high school (omitted)	--	--	--	--
more than high school	-0.005	1.33	-0.005	1.25
Health measures:				
high blood pressure	0.004	1.25	0.003	0.88
diabetes	0.009	2.48	0.007	1.85
cancer/Tumor	0.031	7.85	0.029	7.06
lung disease	0.004	0.96	0.001	0.31
ever heart attack	0.013	3.34	0.012	2.84
angina	-0.009	1.58	-0.012	1.98
congestive heart failure	0.014	2.12	0.011	1.68
stroke	0.013	2.42	0.011	2.08
arthritis/rheumatism	-0.003	1.12	-0.005	1.74
BMI low	0.005	1.07	0.003	0.66
BMI high	0.003	0.63	0.002	0.42
BMI missing	-0.001	0.09	0.000	0.02

Source: Authors' calculations from HRS waves 1 and 2
Number of observations: 10452

Table A-1: Determinants of subjective probabilities (wave 2)

	P75		P85	
	Coeff	Std Err	Coeff	Std Err
Intercept	0.503*	0.053	0.239*	0.057
Household income:				
lowest quartile (omitted)	--	--	--	--
2nd quartile	0.018*	0.009	-0.009	0.010
3rd quartile	0.018	0.010	0.001	0.011
highest quartile	0.016	0.010	-0.002	0.011
Household wealth:				
lowest quartile (omitted)	--	--	--	--
2nd quartile	-0.003	0.009	-0.019	0.010
3rd quartile	-0.011	0.010	-0.023*	0.011
highest quartile	0.017	0.011	-0.008	0.012
Age	0.003*	0.001	0.004*	0.001
Married	-0.012	0.008	-0.010	0.009
Nonwhite	0.059*	0.008	0.112*	0.009
Male	-0.024*	0.007	-0.046*	0.008
Light physical activity:				
3+ times per week (omitted)	--	--	--	--
1-2 per week	-0.018*	0.008	-0.021*	0.009
1-3 per month	-0.025	0.020	-0.041	0.022
< 1 per month	-0.056	0.054	-0.075	0.061
never	-0.014	0.010	-0.009	0.012
Heavy physical activity:				
3+ times per week (omitted)	--	--	--	--
1-2 per week	-0.013	0.009	-0.014	0.010
1-3 per month	-0.020	0.013	-0.023	0.014
< 1 per month	-0.086*	0.025	-0.075*	0.027
never	-0.028*	0.008	-0.033*	0.009
Health status:				
excellent (omitted)	--	--	--	--
very good	-0.055*	0.009	-0.081*	0.010
good	-0.109*	0.009	-0.137*	0.010
fair	-0.204*	0.012	-0.224*	0.014
poor	-0.288*	0.017	-0.258*	0.018

Smoking status:				
never smoked (omitted)	--	--	--	--
former smoker	0.018*	0.007	-.010	0.008
smoker	-0.025*	0.008	-0.017	0.009
Drinking behavior:				
does not drink (omitted)	--	--	--	--
drinks < once/day	0.021*	0.007	0.019*	0.008
1-2 /day	0.019	0.011	0.014	0.012
3-4 /day	-0.012	0.018	-0.002	0.021
drink 5+	0.026	0.030	0.010	0.035
Schooling:	-0.013	0.008	0.012	0.009
less than high school	-0.013	0.008	0.012	0.009
high school (omitted)	-	--	--	--
more than high school	0.015*	0.007	0.020*	0.008
Parental mortality:				
mom alive	0.046*	0.016	0.065*	0.019
mom's age -65 if alive	0.000	0.001	0.001	0.001
mom died age < 51	0.023	0.015	0.039*	0.016
mom died age 51-64	-0.007	0.013	0.010	0.014
mom age at death -65 if > 64	0.002*	0.001	0.004*	0.001
dad alive	0.045*	0.022	0.062*	0.030
dad age-65 if alive	0.001	0.001	0.001	0.002
dad died age < 51	0.032*	0.012	0.055*	0.014
dad died age 51-64	-0.001	0.011	0.011	0.011
dad's age at death - 65 if >64	0.003*	0.001	0.004*	0.001
Diseases:				
high blood pressure	-0.008	0.006	-0.011	0.007
diabetes	-0.050*	0.010	-0.030*	0.011
cancer/Tumor	-0.031*	0.012	-0.019	0.014
lung disease	-0.022*	0.011	-0.034*	0.012
heart attack	-0.029*	0.010	-0.039*	0.012
angina	-0.005	0.017	-0.010	0.019
congestive heart failure	-0.032	0.024	0.027	0.026
stroke	0.005	0.018	0.009	0.020
arthritis/rheumatism	0.000	0.006	-0.002	0.007
Mean of dependent variable	0.637		0.415	
Number of observations	7730		7699	