Probabilistic Thinking and Early Social Security Claiming

Adeline Delavande
RAND Corporation, Universidade Nova de Lisboa and CEPR

Michael Perry
University of Michigan

Robert J. Willis
University of Michigan

Updated for the 8th Annual Joint Conference of the Retirement Research Consortium
"Pathways to a Secure Retirement"
August 10-11, 2006
Washington, D.C.

The research reported herein was pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium (RRC). The findings and conclusions expressed are solely those of the authors and do not represent the views of SSA, any agency of the Federal Government or the RRC. Willis acknowledges support from NIA grant P01-AG026571. Finally, the Health and Retirement Study data used in this paper is supported by NIA U01-AG009740.

1. Introduction

Standard economic theory predicts that individuals decide to delay claiming of their Social Security benefits if doing so increases their expected lifetime utility. In particular, Life Cycle Models predict that those who expect to be long-lived will delay application for benefits because they perceive the increase in Social Security benefits resulting from claiming later as financially beneficial.

By design, Social Security benefits are intended to be actuarially neutral regardless of when an individual retires between the early claiming age of 62 and the upper limit of age 70. That is, the level of monthly benefits for a person who delays claiming by a given amount of time is increased sufficiently to make the expected present discounted value of the annuity stream that he or she will receive invariant to the age of retirement. This design is intended to allow individuals to exercise their preferences for early or late retirement without either increasing or decreasing the cost borne by taxpayers in financing the Social Security system.

The actuarial adjustment of benefits cannot be truly neutral for the entire eligible population (of people with the same earnings histories), if there is individual-level heterogeneity in mortality rates. Individuals who face mortality rates that are higher than actuarial rates will, on average, receive lower lifetime benefits than the representative individual who retires at the same age and, conversely, those with low mortality risk will tend to receive high lifetime benefits. These differentials increase horizontal inequality in lifetime Social Security benefits, but have no effect on the neutrality of the actuarial adjustment from the standpoint of taxpayers unless people act on these differences by choosing different claiming ages, thus creating losses due to adverse selection from the taxpayers' point of view. That is, persons with high mortality rates mitigate their lower benefits by claiming early and persons with low mortality maximize their advantage by claiming late, with both types of behavioral response increasing the expected cost to taxpayers.

In a recent paper, Hurd, Smith and Zissimoupolos (2004), denoted HSZ, attempted to determine this behavioral effect by estimating the effect of subjective survival beliefs on the probability of Social Security claiming and retirement using data from the Health and Retirement Study (HRS). The advantage of using subjective survival responses is that variables determining health and wealth tend to be correlated with mortality expectations, making it difficult in analyses using conventional economic variables to identify effects due to expectations alone. HSZ argue that the direct measures of survival expectations contained in the Health and Retirement Study allow for identification of mortality expectations because, despite being correlated with health and wealth, there is much individual variation in expectations which permits identification of the effect of mortality expectations on claiming and retirement behavior. In their empirical results, they find that "…subjective survival of 0 is significantly associated with higher levels of both retiring and of claiming SS benefits" but no pattern is found for other levels of beliefs.

We are returning to the same question because we believe that the effect they estimated might be too small, primarily because of measurement error in survey answers to subjective probability questions which cause coefficient inconsistency. In addition, we examine claiming over a longer period than HSZ. Like HSZ, we find the survey response to the survival question does not have a statistically significant effect on claiming behavior and the point estimate is near zero. However, when we use instrumental variable methods to correct for measurement error, we find that, among people who are still working at age 62, those who expect to live longer are likely to delay claiming of Social Security benefits to a degree that is both statistically and economically significant. For example, using the estimated claiming equation in a bivariate probit model of retirement and claiming, an increase of 5 percentage points in the predicted survival probability of each person (with an upper limit of 100 percent) leads to a 1.9 percentage point decline in the proportion who claim before age 64, from 29.6 percent to 27.7 percent. For people who are retired by age 62, we find no effect of subjective survival on claiming, even after using instrumental variables.

Our results suggest that subjective probabilities elicited on a survey can be useful in understanding heterogeneity in decisions and behavior, albeit with the issue of measurement error requiring attention. Without direct measurement of subjective probabilities, Manski (2004) argues that economic models of behavior under uncertainty simply make a priori assumptions about people's probability beliefs, typically assuming that people have "rational expectations" and relying on revealed preference to account for behavior. In the context of mortality, the assumption of rational expectations is often taken to mean that an individual's survival probabilities are equal to those found in standard actuarial tables. It is clear, however, that individuals commonly have more information about their health, family history and other factors that might influence their mortality expectations than is available in standard life tables which only distinguish age, sex and sometimes race. In a longitudinal survey such as the HRS which follows individuals over a considerable length of time and has good data on their mortality, it is possible to implement much richer versions of the rational expectations approach in which the econometrician estimates a model of mortality risk based on the relationship between actual mortality and a large set of risk factors measured in the survey.

In this paper, we explore the actuarial approach by estimating a prediction equation in which baseline information from the HRS is used to predict whether an individual is still alive eight to twelve years later with the same covariates used to instrument subjective probability beliefs described above. The correlation between predicted values of actual survival and subjective survival is about 0.5, indicating that the weighting of risk factors by an actuary looking at objective outcomes is substantially different than the weighting of the same factors given by a survey respondent in formulating his or her subjective probability beliefs. Despite these differences, if we enter the actuarial prediction into our bivariate probit model of retirement and claiming, the coefficient on survival probability is close in magnitude, but somewhat less precisely estimated, than the coefficient of the predicted subjective survival probability variable in our first model. We also estimate a model in which we use the predicted subjective survival variable plus another variable measuring the difference between the predicted objective and subjective survival measures to see if claiming behavior is sensitive to any of the measured risk factors beyond their effect on subjective probability beliefs. This latter variable is insignificant and the coefficient of predicted subjective survival remains essentially unchanged in magnitude and significance.

Our results suggest that subjective survival probabilities can be used to capture meaningful behavioral responses to incentives for early claiming in Social Security when they are purged of measurement error using risk factors as instruments. In addition, they are consistent with a rational expectations perspective in the sense that we obtain the same behavioral response if the weights of risk factors are determined by their ability to explain actual mortality and, finally, the risk factors do not explain any claiming behavior beyond their effect on subjective probability beliefs. It is important to point out, as Delavande and Rohwedder (2006) stress, that subjective probability measures can be obtained cheaply in cross-section data sets on a given cohort, allowing researchers to study the effects of mortality on behavior immediately without waiting for the lengthy period of time necessary to observe significant mortality in the cohort.

2. The Data

2.1. The HRS

The HRS is a nationally representative panel survey of persons born in 1953 or earlier, designed to investigate retirement behavior and its implications on the health, social, and economic status of the aging population in the US. Various cohorts were enrolled at different points in time and were interviewed every two years after enrollment.[1] Spouses of age-eligible respondents were also interviewed regardless of age.[2] The present paper uses respondents from the original HRS cohort of those born in 1931 through 1941 and the AHEAD cohort of those born between 1890 and 1923, as well as their spouses.

The HRS has collected subjective probabilities of survival until a target age since its baseline in 1992 using the following question: "***What is the percent chance that you will live to be age [X] or more?***", where the target age *X* depends on respondents' age at the time of interview. Those questions have been the object of several validation studies focusing on their accuracy. Hurd and McGarry (1995) find that average survival probabilities are very close to those presented in life tables and co-vary with variables such as smoking, drinking, health conditions or education in ways that would be expected from studies of actual mortality. In a more recent paper, Hurd and McGarry (2002) use panel data from HRS and find that respondents modify their probabilities in response to new information such as the onset of a new illness.

---

[1] As of 2004, the HRS combines the following cohorts: (a) the original HRS cohort of those born in 1931 through 1941 and studied since 1992; (b) the AHEAD cohort of those born between 1890 and 1923 and studied since 1993; (c) the Children of the Depression Age cohort of those born between 1924 and 1930 and (d) the War Babies cohort of those born between 1942 and 1947, both added to the HRS in 1998; and (e) the Early Baby Boomers cohort of those born between 1948 and 1953, who were added in 2004.

[2] Source: HRS website at http://hrsonline.isr.umich.edu and St Clair et al, 2006.

Figure 1: Percentage of Social Security claiming by respondents who retired before age 62 and claim by 2004 (N=1,801)[3]

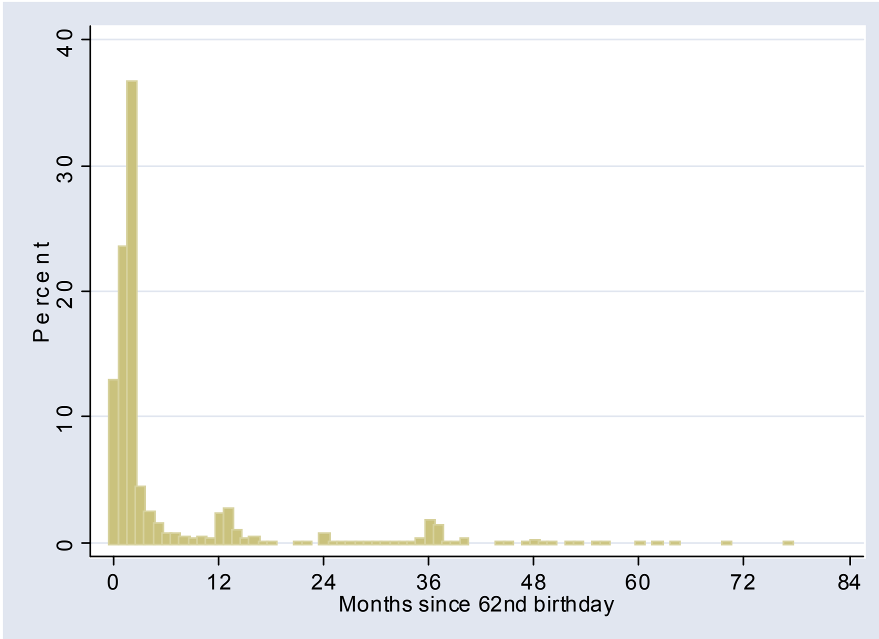
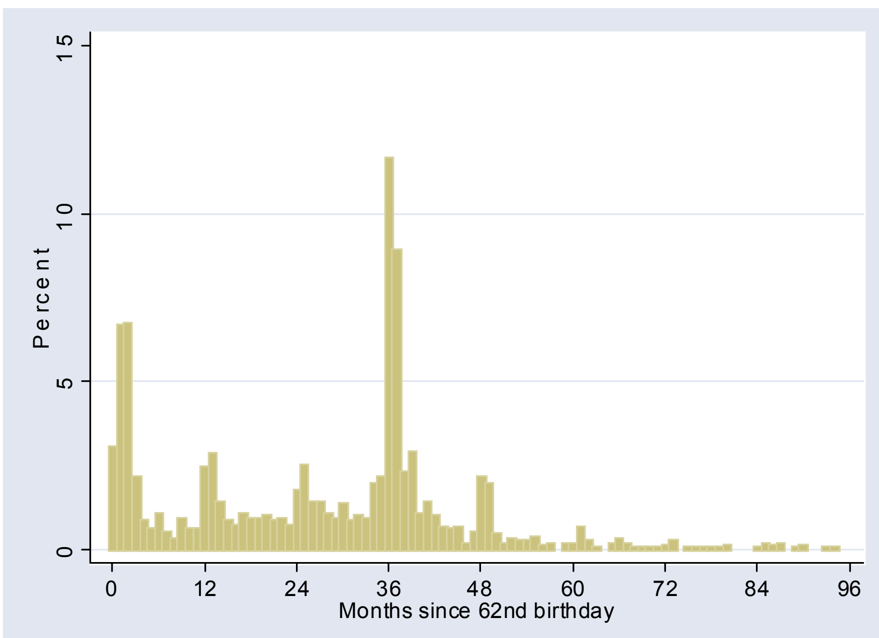


Figure 2: Percentage of Social Security claiming by respondents who are not retired by age 62 and claim by 2004 (N=1,473)



[3] The graph excludes one respondent who reports claiming 153 months after age 62.

2.2. The Analytical Samples

We follow HSZ's approach and use two groups of HRS respondents in our analysis. First, we examine the group that consists of all respondents who retire prior to their 62nd birthday (group 1). Second, we examine the group that consists of all respondents who retire after turning 62 (group 2).[4] We focus on age 62 because workers eligible to claim Social Security benefits may do so at the Early Retirement Age of 62 years. Benefits are permanently reduced if the worker claims before the Normal Retirement Age (which ranges between 65 and 66 for the respondents in our samples) and there is delayed retirement credit which increases permanently a worker's benefit if she claims after her Normal Retirement Age.[5] We conduct separate analysis for each of these two groups as the claiming decision of respondents who are not retired at age 62 is likely to be related to their retirement decision.

We pool respondents across waves and examine the first group at the time they were either 61 or 62 depending on when they were interviewed (60% are age 62). Similarly, for the second group, we look at each respondent at the wave in which they were either 62 or 63 (53% are 62). For each group we model the decision to claim social security, and for the second group we also model the decision to retire, although this decision is not our primary focus. As in HSZ, we exclude individuals who claim Disability Insurance, are widows and receive Social Security benefits before age 62.

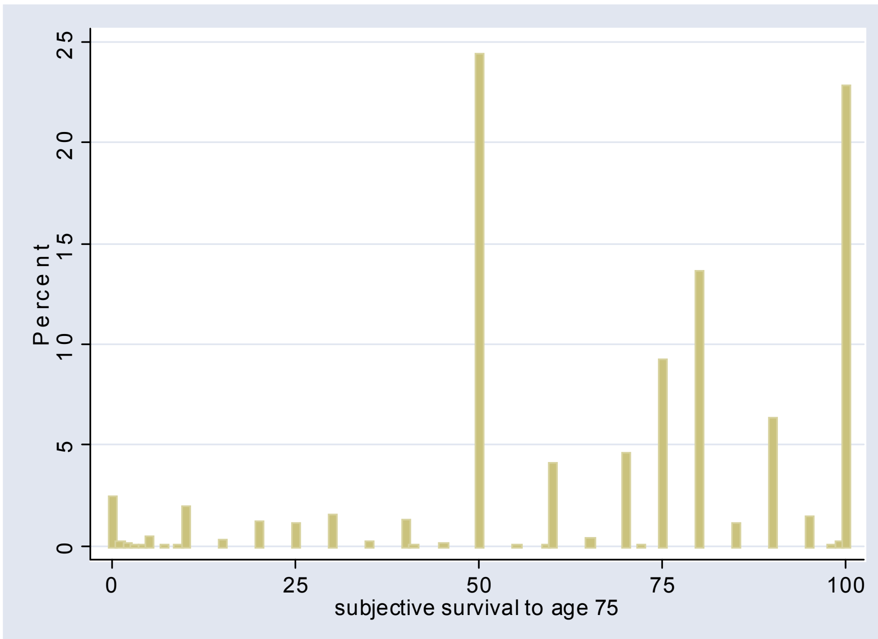
Table 1 presents the distribution of age of first claiming social security for the two groups. The group who retire before 62 claim on average two years earlier than those who retire later (age 62 and 1 month for the early retirement group and age 64 and 1 month for the later retirement group). Figure 1 presents the percentage of Social Security claiming by respondents who retired before age 62 and illustrates that few people in this group delay Social Security claiming: 79.2% claim by the first year after turning 62, 85.5% by the second year and 89.6% by the 3rd year. In contrast, the distribution of claiming for respondents who are not yet retired at 62 is flatter, with the largest spike at 65 (Figure 2). Among this group, only 21.2% claim by the first year, 35.3% by the

[4] We consider a respondent retired if she reports to be fully or partially retired.

[5] The delayed retirement credit per year ranges from 4.0% to 7.5% for the respondents in our sample, and the reductions of benefits for early claiming are or 5/9 of 1% per month for the first 36 months and 5/12 of 1% for subsequent months

second year and 62.2% by the 3rd year. In our analysis, the main variable that we model is whether the respondent claims by age 64 (which corresponds to 24 months in the Figures).

Figure 3: Distribution of subjective probability of survival until age 75 for respondents who are not retired by age 62 and claim by 2004 (N=1,615)



Our analysis uses extensively respondents' answers to their subjective survival until age 75. Figure 3 shows the distribution of subjective survival responses for respondents who are not retired by 62, and Table 2 presents summary statistics for each group's distributions. Most of the respondents provide a large probability: 89% give an answer larger than 50%, the most common answers being 50, 100 and 80. The distributions of subjective survival responses are virtually identical for the two groups, although the early retirement group has a slightly lower mean (67% versus 69%). The median subjective belief of survival to age 75 is 75% in both groups.

3. The effects of survival expectations on claiming behavior

This section presents our analysis of the effect of survival expectations on claiming behavior for the two groups described in 2.2. For the early retirement group, we focus solely on claiming behavior. For the later retirement group we analyze the joint decision to retire and claim Social Security benefits. For each group we use four different specifications: one in which the independent variable of interest is the subjective survival response from the HRS (as in HSZ); one in which we instead used instrumented subjective survival values with subjective survival responses as the independent variable; one in which we used instrumented survival values based on actual mortality in the sample (hereafter referred to as "objective"); and finally one in which we used both the instrumented subjective survival values and the difference between the instrumented subjective and objective survival probabilities.

3.1 The HSZ specification

This subsection follows closely the specification of HSZ to determine the role of subjective survival on claiming. Our analysis for the early retirement group consists of a probit model in which the dependent variable is a dummy indicating the respondent claimed social security before age 64. For the later retirement group we use a bivariate probit model on the joint decision to retire by age 64 and claim social security by age 64.[6,7]

In addition to the subjective survival response, we use a set of covariates to account for other factors that may influence the decision to claim social security and/or retire, such as education, race, gender and wealth. We use different measures of wealth: (1) Financial wealth; we define the wealth quartile for non-housing financial wealth where the quartiles are defined separately for each HRS birth cohort and marital status (single/couple); (2) Individual Retirement Account wealth quartile and (3) Social Security wealth quartile. Note that we do not use administrative data for obtaining Social Security wealth. We rely instead on the reported Social Security benefits for respondents who have claimed before the last wave of HRS in 2004. For the censored observations (i.e. respondents who have not yet claimed by 2004), we use the reported expected Social Security benefits elicited from respondents. The distributions of actual and expected Social Security benefits are very similar.[8] Note however that two individuals with the same earnings history could have different observed (or expected) benefits because they have claimed at a different age. To make the measure of Social Security wealth

[6] Note that 64 was not the threshold used by HSZ. Due to the shorter length of their panel, they used 62 + 2 months for the early retirement group and 63 years old for the later retirement group. We take advantage of new HRS waves to look at a longer delay in claiming.

[7] Our results are not sensitive to the choice of age 64. We obtain very similar results using the same specifications for claiming and retirement by ages 63 and 65.

[8] For example, for group 2, the 25th, 50th and 75th percentiles of realized annual SS benefits are 6960, 10800 and 14400 respectively while those of the expected benefits are 7200, 10800 and 13200.

comparable across people who have claimed at different ages, we compute what would have been the respondent's benefits if she had claimed at 65 using the delayed retirement credit and the reduction of benefits provided by the Social Security Administration for each of the birth years of our respondents (from 1928 to 1942). For this computation, we use the actual claiming age or expected claiming age for respondents who have not yet claimed. For the group of respondents who are not retired by age 62, we also include variables about their current job, such as wage and whether the job requires physical activity.

In this specification we find essentially the same result as in HSZ: subjective survival responses have essentially no effect on claiming. The first column of Tables 3 and 4 shows the results of the probit for the early retirement group and the bivariate probit for the later retirement group respectively. Estimated coefficients for both groups for subjective survival are very close to zero and have very high p-values.

Males and more educated respondents are more likely to delay claiming. High financial wealth should reduce early claiming by easing liquidity constraints. However, financial wealth is not significant in these regressions. Workers with more Social Security wealth are more likely to claim early.

3.2 Correcting for Measurement Error in the Subjective Probability of Survival

3.2.1. Instrumental Variables Approach

As pointed out in Section 2, individual subjective survival probabilities behave as expected when averaged across respondents but they appear to contain considerable noise at the individual level. Rounding to the nearest 5 and heaping at "focal" values of 0, 50 and 100 are common. For example, in each of the sample that we use for analysis, almost half of the respondents report either 50 or 100 when asked about their chance of surviving to age 75. To use the subjective survival responses to make credible inference on individual decision-making, it is thus crucial to correct for measurement error.

We use instrumental variable methods to correct for measurement error. The instruments include 4 sets of variables: (i) some basic demographic characteristics (gender, race, marital status and number of children), (ii) an array of health variables (self-reported health and whether the respondent was diagnosed with some conditions such as diabetes and cancer), (iii) dummy variables on parental mortality (own and spouse) and (iv) an optimism index.

Gender, race and marital status are known to influence life expectancy (CDC 2003, Oswald and Gardner 2002). Parental mortality captures genetic factors that are widely recognized to be important in determining life expectancy. Moreover, parental mortality has been found to influence individual mortality expectations in earlier studies (Hammermesh 1985, Hurd and McGarry 1995, 2002). For example, using data from the HRS, Hurd and McGarry (2002) show that the death of a parent is associated with a reduction of subjective survival across waves. We also include parental mortality of the spouse because observing the death of someone in one's entourage might influence mortality expectations.[9] Finally, we include an optimism index. The optimism index is a predicted value for each respondent based on factor analysis of other probability responses (not including the subjective survival responses). It is calculated separately in each wave because each wave contains different questions about events that most people would consider clearly good or bad. In other analyses (Kezdi and Willis 2002; Hill, Perry and Willis 2004) this index has been shown to be correlated with subjective survival responses and behavior such as stock holding. More information about the optimism index is available in Kezdi and Willis (2002).

Table 5, column I, presents the first stage regression of the instrumental variables model for the group who are not retired prior to their 62nd birthday; i.e. the coefficients obtained from the linear regression of our set of instruments on the elicited probability of surviving until age 75.[10] Most of the coefficients are in the expected direction, e.g., being a male and having a lung disease is associated with a lower subjective probability of survival. Self-reported health (both excellent/good and fair/poor) has the largest effect on beliefs, followed by other health variables and parental mortality. Being black is also associated with a higher probability of survival and the coefficient is both large and significant. This goes against the empirical finding that blacks have lower life expectancy. The optimism index is significant and has a relatively large effect. We do not find that the mortality of the spouse's parents influence mortality expectations.

---

[9] Tversky and Kahneman (1974) find that people tend to use the availability heuristics when judging the likelihood of events, i.e. they tend to be biased by information which is easy to recall.

[10] The first stage for the early retirement group is basically identical and is not presented to conserve space.

3.2.2. The effects of instrumented survival expectations on claiming behavior

The third and fourth columns of Tables 3 and 4b presents the probit and bivariate probit results for each of the groups that relate survival expectations to claiming. We now use the instrumented subjective survival to age 75 and all the other covariates as described in 3.1. For the early retirement group correcting for measurement error did not change our results: the instrumented subjective survival is still not significant[11].

For the later retirement group the estimated coefficient in the bivariate probit is -0.016 (p-value=0.004). The predicted mean early claiming probability in the population of late-retirees is 29.6% which corresponds to a mean subjective survival probability (instrumented) of 69%. Increasing the survival probability to 74% (by shifting everyone up five percentage points and capping them at 100) decreases predicted early claiming to 27.7%. Over a larger range, as survival probability varied from 59% to 79%, early claiming probability varied from 34% to 26%.

That we find an effect of subjective survival for late-retirees and no effect for early-retirees is not too surprising due to the difference in claiming behavior already shown in figures 1 and 2. That is, the vast majority of early-retirees claim very early, so there is little variation in claiming behavior to be explained.

Also notable is that in the bivariate results for the later retirement group, rho, the correlation between the regression errors in the retirement specification and the claiming specification, is very large and positive. This indicates that many of the same unobservables drive both decisions.

One might be concerned about the inclusion of parental vital status in our set of instruments because having an elderly parent might influence labor supply decision, especially of women (it is not clear however how it may influence claiming). To rule out the fact that having a mother or father alive influences the retirement decision directly in addition to its effect on expectations, we include the dummy for mother and parent alive in the retirement equation of the bivariate probit model (both the instrument and non-

[11] Standard errors for all specifications that include instrumented first-stage variables as explanatory variables were calculated using 1000 bootstrap repetitions of both the instrumenting and probit or biprobit specifications. Bootstrapping was necessary because standard techniques produce incorrect standard errors when predicted values from an instrumental variables specifcation are included as independent variables.

instrument specifications). Those dummy variables do not statistically significantly influence the retirement decision (tables not shown).

3.3. Subjective and Objective Survival Compared

3.3.1. Actuarial Estimates

As pointed out in the introduction, researchers who do not have data on subjective probabilities and seek to conduct inference on behavior under uncertainty need to make unverifiable assumptions about people's expectations. In this section, we take advantage of the longitudinal aspect of the HRS to investigate the results one would get following this approach. The HRS provides detailed information on vital status of respondents through tracking and matches with the National Death Index.[12] We construct the "objective" risk of survival to age 75 using information on the actual vital status of the respondents in a later wave.

More specifically, we estimate a logit model using being alive in 2004 as the dependent variable and the same set of covariates as the one used to instrument the subjective survival response described in Section 3.2.1. We estimate a separate model for each group, and restrict our analysis to respondents who were 61 and 62 (62 and 63 respectively) before 1994 for the early retirement group (the later retirement group, respectively). The motivation for selecting those respondents is that they are either 75 or just a few years younger than 75 in 2004 when we observe their vital status. Table 5, column II, presents the coefficients of the logit estimate for the respondents who are retired by 62. Among those, 18% of the respondents have died by 2004.[13]

While the coefficients cannot be compared directly with the ones from the first column of Table 5, the results indicate that the weighting of risk factors by an actuary looking at objective outcomes is substantially different than the weighting of the same factors given by a survey respondent in formulating his or her subjective probability

[12] The respondent is considered alive by HRS if she was interviewed, or contacted directly by an interviewer during the wave, or was said to be alive by a spouse or partner, or was not reported dead (source: Data description and usage of the tracker file 2002). In addition, for up to 2002, the HRS seeks matches to the NDI for respondents who are reported deceased or who are of unknown vital status during tracking.

[13] The logit results for the early retirement group are similar.

beliefs. For example, while being black is associated with a higher subjective probability of survival, it is associated with a lower objective probability survival. Moreover, the correlation between predicted values of actual survival and subjective survival is 0.68 for the early retirement group and 0.45 for the later retirement group, indicating substantial differences in the measures.

We use the coefficients presented in the second column of Table 5 to impute an objective probability of survival to age 75 for all the respondents in the later retirement group, and follow a similar procedure for the early retirement group.

3.3.2. The Effects of Objective Survival on Claiming Behavior

The fifth and sixth columns of Tables 3 and 4b show the estimated effect of our objective probability measure on claiming. For the early-retirement group, the effect remains essentially zero. For the later-retirement group, the effect of the objective probability measure on claiming is very similar to the effect of subjective probability and also measured precisely (-0.013 with p-value 0.007). This suggests that in this application, objective and subjective probability seem to provide about as much information as each other to explain claiming behavior. That is, an analyst using the rational expectations approach in which subjective survival is inferred from actual future mortality would not do any worse than one who had real subjective survival responses.

The last two columns of Tables 3 and 4b show the estimated effect of instrumented subjective survival when the difference between subjective and objective survival is also included as a covariate. Again, for the early-retirement group, we find no effect. For the later-retirement group, the size and precision of the subjective probability coefficient remains very similar to the specification when the difference was not included (-0.019, p-value of 0.003). The coefficient estimated for the difference itself is of fairly small magnitude and is measured with high standard error (p-value 0.234). We interpret this to mean that once the subjective probability has been taken into account, the objective probability adds no power to explain claiming behavior.

4. Conclusion

We present evidence in this paper that people's personal probability beliefs about their survival chances have a statistically and economically significant effect on early Social Security benefit claiming. The paper contains two important methodological contributions. First, we show that the effects of subjective probability beliefs on behavior are almost completely obscured unless measurement error is taken into account. When risk factors measuring features of an individual's demographic characteristics, health and family history are used as instruments to purge measurement error from answers to HRS subjective survival probability questions, we find that an increase of 5 percentage points in the predicted survival probability to age 75 leads to a 1.9 percentage point decrease in the proportion of persons still working at age 62 who will claim benefits before age 64. Second, we show that the estimated effect of subjective probability beliefs on claiming is close in magnitude to the effect of predicted mortality risk based on an actuarial model relating risk factors of HRS respondents to their survival 8-12 years into the future. We also find that beyond their effect on subjective survival beliefs these risk factors have no additional effect on claiming behavior. These findings suggest that people may incorporate mortality risks into their decision making about Social Security claiming in a way that is consistent with rational expectations.

Despite the substantial and growing literature on survey measures of subjective probabilities, there are few instances of research that has attempted to use these variables to explain choice behavior. In that context, the paper represents an early contribution to what we expect to be an important line of research which attempts to understand the relationship between probabilistic beliefs and economic decisions.

From a policy perspective, our results imply that it would be worthwhile to do further research on the role of heterogeneous mortality risk on claiming behavior as well as other decisions about savings, insurance and annuitization where mortality risk is expected on theoretical grounds to play an important role.

References

Center for Disease Control and Prevention. 2003. "United States Life Tables , 2003." National Vital Statistics Report, 54(14).

Delavande, Adeline and Susann Rohwedder. 2006. "Differential Mortality in Europe and the US: Estimates Based on Subjective Probabilities of Survivals." Unpublished manuscript.

Gardner, Jonathan and Andrew Oswald. 2002. "Is it Money or Marriage that Keeps People Alive?" Unpublished manuscript.

Hammermesh, Daniel. 1985. "Expectations, Life Expectancy and Economic Behavior." Quarterly Journal of Economics. 100(2): 389-408.

Hill, Dan, Michael Perry and Robert J. Willis. 2004. "Estimating Knightian Uncertainty From Survival Probability Questions on the HRS." Unpublished manuscript.

Hurd, Michael and Kathleen McGarry. 1995. "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study." *The Journal of Human Resources* 30: S268-S292.

Hurd, Michael and Kathleen McGarry. 2002. "The Predictive Validity of Subjective Probabilities of Survival." *The Economic Journal*. 112: 966-985.

Hurd, Michael, James P. Smith and Julie M. Zissimopoulos. 2004. "The Effects of Subjective Survival On Retirement and Social Security Claiming." *Journal of Applied Econometrics*. 19: 761-775.

Kezdi, Gabor, and Robert J. Willis. 2003. "Who Becomes a Stockholder? Expectations, Subjective Uncertainty and Asset Allocation." Unpublished manuscript.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica*. 72(5): 1329-1376.

St. Clair, Patricia, et. al. 2006. RAND HRS Data Documentation, Version F.

Table 1: Age First Claim Social Security

Group	Mean	5th	25th	50th	75th	90th
Retire after 62	64y, 5m	62y, 1m	63y, 4m	64y, 11m	65y, 2m	67y, 1m
Retire before 62	62y, 5m	62y	62y, 1m	62y, 2m	62y, 3m	65y

Table 2: Subjective Probability of Survival to Age 75

Group	Mean	5th	25th	50th	75th	90th
Retire after 62	69	10	50	75	90	100
Retire before 62	67	10	50	75	90	100

Table 3: Probability of claiming SS by age 64

Probit specification for respondents who are retired by 62

	I		II		III		IV	
	Coef.	*P value*	Coef.	*P value*	Coef.	*P value*	Coef.	*P value*
Subjective Prob.	**0.002**	*0.352*					**0.000**	*0.978*
IV Subjective Prob.			**0.000**	*0.945*				
Objective probability					**0.000**	*0.913*		
Subj - Obj probability							**-0.001**	*0.823*
Male	-0.298	*0.002*	-0.293	*0.006*	-0.288	*0.006*	-0.287	*0.009*
Black	-0.156	*0.283*	-0.200	*0.178*	-0.194	*0.207*	-0.177	*0.327*
Wealth quartile - low	-		-		-		-	
Wealth quartile - 2	-0.024	*0.878*	0.015	*0.928*	0.015	*0.926*	0.014	*0.931*
Wealth quartile - 3	0.083	*0.597*	0.141	*0.405*	0.140	*0.397*	0.143	*0.375*
Wealth Quartile high	0.046	*0.789*	0.083	*0.644*	0.081	*0.657*	0.083	*0.623*
Married	0.280	*0.036*	0.228	*0.126*	0.225	*0.144*	0.222	*0.142*
Own Stock	-0.101	*0.381*	-0.136	*0.251*	-0.136	*0.275*	-0.136	*0.241*
Less than HS	0.026	*0.835*	0.109	*0.418*	0.111	*0.420*	0.109	*0.434*
High School	-		-		-		-	
College or more	-0.523	*0.000*	-0.414	*0.000*	-0.414	*0.000*	-0.415	*0.000*
IRA quartile - low	-		-		-		-	
IRA quartile - 2	-0.016	*0.925*	-0.027	*0.884*	-0.028	*0.880*	-0.025	*0.896*
IRA quartile - 3	-0.108	*0.371*	-0.080	*0.559*	-0.081	*0.563*	-0.081	*0.551*
IRA quartile - high	-0.189	*0.152*	-0.168	*0.250*	-0.169	*0.247*	-0.168	*0.237*
Health Limits Work	0.010	*0.931*	-0.085	*0.504*	-0.075	*0.519*	-0.082	*0.502*
SS quartile - low	-		-		-		-	
SS quartile - 2	0.257	*0.026*	0.308	*0.013*	0.307	*0.015*	0.308	*0.016*
SS quartile - 3	0.925	*0.000*	0.958	*0.000*	0.957	*0.000*	0.957	*0.000*
SS Quartile high	0.902	*0.000*	0.939	*0.000*	0.937	*0.000*	0.939	*0.000*
Constant	0.700	*0.008*	0.085	*0.600*	0.084	*0.603*	0.083	*0.616*
N	1,660		1,530		1,530		1,530	

Table 4a: Probability of retiring by age 64

Bivariate probit specification for respondents who are not retired by 62

	I		II		III		IV	
	Coef.	*P value*	Coef.	*P value*	Coef.	*P value*	Coef.	*P value*
Retire by 64								
Subjective Prob.	**-0.00112**	*0.438*					**-0.011**	*0.159*
IV Subjective Prob.			**-0.009**	*0.132*				
Objective probability					**-0.007**	*0.167*		
Subj - Obj probability							**0.003**	*0.581*
Male	-0.037	*0.669*	-0.015	*0.877*	-0.086	*0.418*	-0.049	*0.664*
Black	-0.053	*0.670*	0.084	*0.542*	-0.035	*0.817*	0.042	*0.789*
Wealth quartile - low	-		-		-		-	
Wealth quartile - 2	-0.340	*0.002*	-0.437	*0.000*	-0.435	*0.001*	-0.436	*0.000*
Wealth quartile - 3	-0.157	*0.156*	-0.193	*0.108*	-0.189	*0.102*	-0.193	*0.107*
Wealth Quartile high	-0.213	*0.107*	-0.205	*0.168*	-0.206	*0.155*	-0.203	*0.153*
Married	0.128	*0.217*	0.125	*0.296*	0.139	*0.226*	0.132	*0.270*
Own Stock	0.100	*0.283*	0.116	*0.267*	0.111	*0.244*	0.114	*0.269*
Less Than HS	0.052	*0.604*	0.101	*0.338*	0.121	*0.284*	0.107	*0.342*
High School	-		-		-		-	
College or more	-0.118	*0.241*	-0.108	*0.350*	-0.117	*0.301*	-0.107	*0.341*
IRA quartile - low	-		-		-		-	
IRA quartile - 2	0.130	*0.652*	0.134	*0.672*	0.134	*0.667*	0.134	*0.669*
IRA quartile - 3	-0.052	*0.584*	-0.032	*0.747*	-0.033	*0.748*	-0.031	*0.758*
IRA quartile - high	0.015	*0.885*	0.035	*0.757*	0.030	*0.793*	0.036	*0.752*
Wage	0.001	*0.523*	0.002	*0.370*	0.002	*0.405*	0.002	*0.383*
Health Limits Work	0.227	*0.122*	0.229	*0.133*	0.251	*0.112*	0.227	*0.169*
Job Req. Physical Act	-0.022	*0.561*	-0.029	*0.491*	-0.029	*0.498*	-0.029	*0.504*
White Collar	-0.259	*0.005*	-0.289	*0.003*	-0.299	*0.003*	-0.293	*0.003*
Pension on Current Job	0.181	*0.018*	0.165	*0.036*	0.167	*0.044*	0.164	*0.052*
Health v. good/excellent	0.003	*0.972*	0.054	*0.538*	0.038	*0.650*	0.055	*0.563*
Health poor/fair	0.089	*0.433*	0.085	*0.502*	0.099	*0.449*	0.084	*0.526*
SS quartile - low	-		-		-		-	
SS quartile - 2	-0.075	*0.497*	-0.113	*0.342*	-0.106	*0.379*	-0.111	*0.348*
SS quartile - 3	0.071	*0.519*	0.055	*0.647*	0.051	*0.661*	0.055	*0.649*
SS Quartile high	0.035	*0.761*	0.022	*0.857*	0.030	*0.803*	0.026	*0.830*
Constant	0.251	*0.313*	0.831	*0.075*	0.831	*0.108*	0.995	*0.123*

Table 4b: Probability of claiming SS by age 64

	I		II		III		IV	
	Coef.	*P value*	Coef.	*P value*	Coef.	*P value*	Coef.	*P value*
Claim by 64 Specification								
Subjective Prob.	**-0.002**	*0.132*					**-0.019**	*0.003*
IV Subjective Prob.			**-0.016**	*0.004*				
Objective probability					**-0.013**	*0.007*		
Subj - Obj probability							**0.006**	*0.234*
Male	-0.372	*0.000*	-0.401	*0.000*	-0.541	*0.000*	-0.471	*0.000*
Black	0.035	*0.786*	0.071	*0.581*	-0.148	*0.340*	-0.015	*0.927*
Wealth quartile - low	-		-		-		-	
Wealth quartile - 2	-0.151	*0.185*	-0.171	*0.172*	-0.166	*0.173*	-0.167	*0.192*
Wealth quartile - 3	-0.028	*0.808*	-0.059	*0.637*	-0.055	*0.663*	-0.059	*0.631*
Wealth Quartile high	-0.218	*0.116*	-0.212	*0.182*	-0.215	*0.172*	-0.208	*0.188*
Married	0.148	*0.169*	0.085	*0.489*	0.109	*0.376*	0.098	*0.414*
Own Stock	-0.053	*0.584*	0.002	*0.985*	-0.008	*0.935*	0.000	*0.997*
Less Than HS	0.055	*0.589*	-0.011	*0.921*	0.029	*0.798*	0.001	*0.992*
High School	-		-		-		-	
College or more	-0.172	*0.118*	-0.153	*0.203*	-0.170	*0.141*	-0.150	*0.209*
IRA quartile - low	-		-		-		-	
IRA quartile - 2	-0.491	*0.123*	-0.356	*0.699*	-0.376	*0.594*	-0.360	*0.723*
IRA quartile - 3	-0.065	*0.508*	-0.073	*0.500*	-0.078	*0.464*	-0.072	*0.489*
IRA quartile - high	-0.164	*0.140*	-0.124	*0.315*	-0.134	*0.254*	-0.124	*0.303*
Wage	-0.002	*0.214*	-0.002	*0.840*	-0.002	*0.848*	-0.002	*0.842*
Health Limits Work	0.084	*0.570*	0.021	*0.899*	0.062	*0.695*	0.015	*0.930*
Job Req. Physical Act	-0.021	*0.596*	-0.025	*0.574*	-0.024	*0.569*	-0.025	*0.571*
White Collar	-0.211	*0.028*	-0.242	*0.019*	-0.264	*0.017*	-0.251	*0.023*
Pension on Current Job	-0.385	*0.000*	-0.378	*0.000*	-0.377	*0.000*	-0.381	*0.000*
SS quartile - low	-		-		-		-	
SS quartile - 2	0.177	*0.120*	0.125	*0.306*	0.138	*0.268*	0.128	*0.307*
SS quartile - 3	0.179	*0.122*	0.195	*0.139*	0.184	*0.154*	0.192	*0.146*
SS Quartile high	0.119	*0.336*	0.157	*0.296*	0.166	*0.256*	0.164	*0.265*
Constant	0.285	*0.270*	0.133	*0.438*	0.113	*0.521*	1.685	*0.003*
Arctan (rho)	0.824	*0.000*	0.798	*0.000*	0.798	*0.000*	0.798	*0.000*
Rho	0.675		0.658		0.659		0.658	
N	1,252		1,159		1,159		1,159	

Table 5: Instrumenting subjective and objective mortality for early-retirees

	Linear regression on the subjective probability of surviving to age 75		Logit regression on being alive in 2004	
	I		II	
	Coef.	*P>t*	Coef.	*P>z*
Male	-1.426	*0.311*	-1.132	*0.000*
Married	1.059	*0.714*	0.275	*0.617*
Black	6.056	*0.006*	-0.669	*0.048*
Number of children	0.164	*0.624*	-0.013	*0.817*
Self-reported health				
very good/excellent	7.113	*0.000*	0.376	*0.173*
poor/fair	-7.367	*0.001*	-0.288	*0.434*
Ever had:				
High blood pressure	-0.215	*0.875*	-0.205	*0.418*
heart problems	-4.785	*0.017*	-0.439	*0.171*
cancer	-2.670	*0.288*	-1.018	*0.024*
diabete	-4.835	*0.027*	-0.557	*0.108*
lung disease	4.589	*0.205*	0.140	*0.860*
stroke	-4.094	*0.363*	-1.390	*0.019*
Parental mortality				
mother alive	0.459	*0.806*	-0.196	*0.580*
mom died before 76	-3.494	*0.042*	-0.047	*0.885*
mom died between 77-84	-		-	
mom died after 85	0.761	*0.741*	0.011	*0.981*
father alive	3.013	*0.300*	-0.780	*0.113*
dad died before 72	-3.861	*0.018*	-0.166	*0.587*
dad died between 73-80	-		-	
dad died after 81	5.220	*0.007*	-0.101	*0.792*
spouse's dad died	0.290	*0.901*	0.086	*0.834*
spouse's mom died	-1.894	*0.232*	-0.157	*0.594*
Optimism index	4.618	*0.000*	0.224	*0.184*
Constant	68.473	*0.000*	2.978	*0.000*
Number of observations	1,439		641	
R2	0.115			
Pseudo-R2			0.095	