Division of Economics
A.J. Palumbo School of Business Administration
Duquesne University
Pittsburgh, Pennsylvania

Forecasting's Final Destination:
How Well Can Individuals Predict their own Death?

Peter J. Horne

Submitted to the Economics Faculty
in Partial Fulfillment of the Requirements for the Degree of
Bachelor of Science in Business Administration

December 2009

Faculty Advisor

Mark T. Gillis December 2009

Faculty Co-Advisor

Dr. Antony Davies December 2009

This paper analyzes individuals' probability expectations of living to the age of 75 using survey data from the Health and Retirement Study. Preliminary specifications find that these expectations are consistent with actual personal mortality risks and also are predictors of an individual's survival. The results also suggest that individuals are, on average, slightly pessimistic in the formulation of such expectations, suggesting the possibility of irrationality in forming longevity expectations. Finally, an examination of the mean deviation in individuals' forecast errors indicates that individuals miscalculate the probability of surviving to age of 75 by an average of about 25 percentage points wherein individuals who survive ultimately estimate more closely than individuals who do not survive.

JEL Classifications: D8

Key words: Mortality Risk, Longevity Expectations, Life Expectancy, Theory of Rational Expectations, Selection Bias, Heckman Two-Step Corrective Procedure, Health and Retirement Study

Table of Contents

I. Introduction

Is an individual capable of predicting his or her own death? Although no one may officially announce such a prediction, the expected duration of one's life is inherently a factor in many decisions that an individual makes in his or her lifetime.[1] Formally, this relates to Friedman's (1957) Permanent Income Hypothesis, and Fisher's (1930), Harrod's (1948), and Ando and Modigliani's (1963) contributions to the Life Cycle Hypothesis in that the foundation of these life cycle consumption and utility models is a rationally conceived longevity expectation. How reliable are such expectations? Are they unbiased? Are they accurate?

At the center of this evaluation is the subjective mortality risks associated with an individual. Mortality risk in a population is regularly calculated actuarially, and then reported in life expectancy tables[2], but it is unlikely that these measures accurately reflect the longevity expectations of each individual. Presumably, individuals have additional information regarding their personal mortality risks and it is therefore reasonable to expect that individuals have a better knowledge of their life expectancy than actuarially calculated life expectancies. However, it is also possible that individuals are systematically upward or downward-biased in their longevity expectations as they react differently to these risks.

The Theory of Rational Expectations, referred to by some authors as simply the "Rationality Hypothesis" or "Assumptions of Rationality," is the notion that individuals consider all relevant information in formulating their forecasts, so that such forecasts remain systematically unbiased in comparison to what is predicted by economic theory (Muth, 1961). Since most Life Cycle and Consumer Choice models presume this rationality assumption,

---

[1] For example, saving for retirement or choosing when to purchase a house.

[2] The U.S. Census Bureau reports life expectancies in their Statistical Abstract. Following is the link to view life expectancy rates for 2009.
http://www.census.gov/compendia/statab/cats/births_deaths_marriages_divorces/life_expectancy.html

establishing the assumption's validity must preface any conclusions made concerning individuals' behavior or plans in future periods. To put it simply, if longevity expectations are in fact rational, then such longevity expectations should be, on average, unbiased. Another important question is even if longevity expectations are unbiased, how accurate are they? Whereas rationality refers simply to being correct on average, accuracy refers to the extent to which individual longevity expectations are actually good predictors of what actually occurs.

Using data from the National Institute on Aging's *Health and Retirement Study*, this paper analyzes individuals' longevity expectations by examining individual predictions of the probability of living to the age of 75. First, this paper examines if known mortality risks are consistent with longevity expectations and if an individual's longevity expectations are reliable and consistent predictors of that same individual's actual mortality. Then, this paper attempts to test whether or not the Theory of Rational Expectations is valid in relation to the individuals' forecasts. Lastly, this paper compares longevity expectations to econometrically estimated probabilities of living to 75 using known information when the longevity expectation was made.

The paper finds results that are consistent with the initial hypotheses that individual longevity expectations, in the form of the predicted probabilities of living to the age of 75, are reflective of subjective mortality risks and are predictors of an individual's survival at the age of 75. However, this paper also suggests that there is potential irrationality in the form of a slight, yet systematic underestimation of individuals' expected chances of living to the age of 75. However, this result depends on whether or not the appropriate model for testing the theory of rational expectations should include future health shocks. Finally, this paper suggests that individuals' predicted probabilities of living to the age of 75 are "off" by about 25% on average, but individuals that are ultimately alive at 75 are more accurate than individuals that are not.

II. Literature Review

The basis for research predicated on evaluating individuals' longevity expectations concerns examining two fields of previous literature. The first field is that of mortality risk. How often people live with or incur risks that are potentially deadly contributes much to their life expectancy. Therefore, economic research which measures these mortality risks and estimates their predictive power on one's probability of survival is relevant to the construction of models that examine the forbearers of one's death. Along with the empirical studies that measure mortality risk, the second relevant topic is that of expectations. In the following sections, each of these fields is reviewed, and then finally, further literature is discussed in which mortality risks and expectations are applied specifically to longevity studies.

A. Mortality Risks

"The moment I climb out of bed I start taking risks," acquiesces Richard Wilson (1979, p. 1). His examination of the mundane risks to life illustrates two tendencies that relate to a discussion of mortality risks. The first is that as time passes and technology progresses, societies are able to mitigate pressing risks. For instance, new medicines, safety features, and surgical procedures are all examples. Then, the second tendency is, given the alleviation of past major risks, society becomes more conscious of other risks; most of which were once deemed minor. Wilson offers statistics on numerous causes of death in the United States, and adds commentary on the willingness of society to accept life-threatening risk on a small scale.[3]

Arthur (1981) defines a more formal framework to understanding mortality risks. He constructs a neoclassical model under the assumptions that individuals leave no bequests, they do

[3] For instance, Wilson acknowledges the following activities that will increase one's risk of death by one in one million: living 2 days in NY or Boston, traveling 6 minutes in a canoe, traveling 300 miles by car, living 2 months in Denver, eating 40 tablespoons of peanut butter, and eating 100 charcoal-broiled steaks.

not fear death, and they seek to maximize their expected lifetime utility. He finds that one's assessment of mortality risks largely depends on one's age. Thaler and Rosen (1976) and Viscusi (1979) examine this relationship empirically in the market for labor by implementing two different interaction terms in two different models: a worker's age multiplied by mortality risk and worker's age multiplied by a worker's expected years of life lost. Then, through a standard hedonic wage equation[4], Thaler and Rosen and Viscusi find a significant and negative age-risk interaction. They conclude that, on average, younger workers face greater mortality risks and uncertainty, and as hypothesized, predict lower probabilities of living to a particular age. Arthur also uses his theoretical model to illustrate the hypothesis that the valuation of one's mortality risk combined with an individual's current age is highly connected to one's longevity expectation. Moreover, Arthur states that changes to one's exposure to mortality risks can be measured by the difference it reflects in one's expected length of life, or work life, as in Thaler and Rosen's or Viscusi's work.

Viscusi, Hakes, and Carlin (1997) further build the connection between mortality risks and longevity expectations by studying the rate at which individuals discount years of life lost in their life expectancies. The researchers develop their study in three stages. The first section accounts for variation in individuals' ages. The second incorporates an individual's subjective discounting rate through including an individual's expected duration of remaining life (a longevity expectation). Thirdly, Viscusi, et al. employ panel data[5] from the U.S. *Vital Statistics* page to introduce lagged exposure to risk and account for health shocks. The researchers find that society discounts years of life lost within rates of 3.3 to 12.4 percent. This variation shows

[4] See Lucas (1977) for a full discussion of the original equation
[5] *Expectations of Life at Single Years of Age, by Race and Sex*

the role of subjectivity in discounting in the presence of mortality risks, and consequently, forming one's longevity expectation.

Hurd, McFadden, and Merrill (1999) expand upon the notions of subjectivity and mortality risks, but hold age relatively constant for their studied cohort, which surveys the predictors of mortality amongst the elderly. Hurd, et al. provide a solid framework for which to build an analysis that evaluates how individuals' longevity expectations may predict death. The researchers include a host of predictors to regress on the mortality of their aged 70 and older Asset and Health Dynamics among the Oldest-Old Study (AHEAD) data, which includes: income levels, education, marital status, and history of heart attack, hypertension, depression, and cancer. Hurd, et al. find that even though the explanatory power of socio-economic variables and mortality declines with age (incriminating any such 'money can buy you health' hypothesis), the host of health indicators remain significant as predictors of mortality.

B. Expectations

The breadth of research concerning the modeling of expectations has been astonishingly prolific since Muth's (1961) seminal paper. However, here only the Theory of Rational Expectations will be outlined with a particular methodological focus targeted in the next section.

Returning to Muth's (1961) original premise, the Theory of Rational Expectations is predicated on an individual's use of information. Muth asserts that individuals, and surely, the economy as a whole, do not waste information. Moreover, and as pertaining to a multi-period model, this entails that individuals do not continually form expectations that consistently depart from what economic theory would predict, and that the only errors in the long run are a product of random deviations. Muth constructed this framework within the microeconomic context of speculation within a restricted hog market in partial equilibrium. Although Muth's analysis was of limited scope, it served three purposes. First, it offered theoretical support for previous investigations concerning price predictions within isolated markets.[6] Second, it introduced the notion that individuals' expectations are not systematically biased; implying that if all relevant information is incorporated, individuals' expectations will only deviate stochastically from economic theory, i.e. the variable's actual value will be equal to expected value of a variable plus a randomly distributed error term:

$$X = E[X] + \varepsilon \tag{1}$$

Third, Muth's work provided the microeconomic foundation for later extensions of broader macroeconomic models.

Lucas (1972) sets forth one of the most noteworthy macroeconomic extensions of Muth's (1961) Theory of Rational Expectations. Lucas constructs an economy comprised of two identical states that exist over two periods, but with no wealth transfers. He then models individuals' consumption and labor preferences and firms' production choices with the government's role being only that of issuing fiat money.[7] This fiat money now introduces the concept of expectations to the consumers. Consumers in each part of the economy must anticipate and then account for the issuance of fiat money in order to consider all relevant information. Moreover, Lucas presumes that individuals accomplish this well, stating that individuals will not be mislead in the long-run by a veil[8] of fluctuations only associated with nominal variables, which, if true, would render monetary policy impotent. However, for the

[6] Heady and Kaldor (1954) study price expectations of eggs, butterfat, cattle, oats, soybeans, corn and hogs through a survey of Iowa farmers between 1947 and 1949. Modligliani and Weingartner (1958) study price expectations as they relate to forecasting sales and designing an investment plan through employing survey data from the Annual Survey of Business Anticipations of Plant and Equipment Expenditures

[7] Fiat Money – Lucas' governments' lever for expansionary policies. This money increases the money supply, but is simply issued – it does not enter the economy through open market operations

[8] John Gurley (1961) referred to money's impact on the real variables of an economy as a veil. The term is synonymous with money illusion or the nominal value of money, instead of its real purchasing power.

context of this paper, the critical aspect of Lucas' theory is the accuracy of individuals' predictions in the presence of anticipated shocks. For if all relevant information and anticipated shocks are incorporated into individuals' expectations, then, individuals will rationally be able to predict the real value of each variable.[9]

Notwithstanding the significance of Muth's and Lucas' research, there still exist Neo-Keynesian empirical studies which present counter hypotheses to the Theory of Rational Expectations. For the purposes of this paper, the review of opposing research concerns the view of known and considered information and predictable shocks as determinants of deviations from a variable's real value.

Mishkin (1982a) tests the Theory of Rational Expectations through the examination of the following questions: Do individuals anticipate changes in monetary policy? Or can the implementation of these policies evoke different responses than an individual's expected response? Moreover, in considering this information, do individuals consistently anticipate, respond, and then act erroneously in facing these policy changes ex ante and ex post? Through employing multivariate time-series specifications, Mishkin specifies several models in which a host of variables[10] is regressed on money growth (M1), GNP, and the unemployment rate. His results show government expenditure and monetary policy variables to be significant over lags up to 20 periods, rejecting the hypothesis that individuals consistently formulate expectations of real variables correctly in the long-run. Moreover, Mishkin finds that when examining the magnitude of the significant coefficients, anticipated movements in monetary policy have larger

[9] Barro (1977), Barro (1978), Barro (1979), Germany and Srivastava (1979) and Small (1979) all test the neutrality implications of monetary policy on macroeconomic variables; most particularly real growth of GDP and all the researchers find evidence which supports the validity of the Theory of Rational Expectations.

[10] Including M1growth lagged the past 4 periods, M2 growth rates, inflation rates, real growth rate of GNP, unemployment rate, t-bill rates, the growth rate of real government expenditure, and the growth rate of government debt – the model's foundations were based off of specifications of Fair (1978)

impacts than unanticipated movements, which directly contradicts the theory that only unanticipated shocks can cause an individual's expectation to differ from the corresponding variable's true value. However, upon an alternate specification that uses output as the dependent variable and inflation and nominal GNP growth as proxies for aggregate demand policy levers, Mishkin finds results favorable to a rationality hypothesis. He notes that unanticipated money growth (a measure for unanticipated shocks)[11] is significant and has the expected signs which could indicate that individuals at least predict in the correct direction. However, upon jointly testing rationality and neutrality hypotheses, a failure of the Granger Causality Test makes these potential implications uncertain. Notwithstanding this obscurity, Mishkin's research provides the foundation for a theory wherein Rational Expectations hold for individuals' predictions, although only directionally, not quantitatively.

To substantiate his previous work to other researchers, Mishkin (1982b) provides a sequel to his investigation of the significance of anticipated shocks on monetary variables and hence rationality assumptions. Here, Mishkin reforms some of his previous specifications. However, the thrust of his analysis remains much the same still regressing monetary policy levers[12] and unanticipated aggregate demand growth variables[13] on real money growth (M1), real GNP, and the unemployment rate. Once again, Mishkin observes that anticipated policy variables produce more significant impacts than unanticipated shocks on all three dependent variables. However, yet again, Mishkin notes that "there is one rejection at the 5 percent level, but just barely" when he tests the rationality hypothesis (Mishkin, 1982b, p. 48). Further, when

[11] Where unanticipated money growth is defined as $M1G_t - M1G_t^e$ also referred to as 'money surprise.'

[12] Quarterly rate of growth of real federal government expenditure, average balance of payments on current account, and the quarterly rate of growth of government debt.

[13] Money surprise, quarterly rate of growth of real GNP, and unanticipated growth in nominal GNP , which is defined as $NGNP_{t-i} - NGNP_{t-i}^e$.

Mishkin jointly tests neutrality and rationality, he states the rejections are primarily due to neutrality postulates as opposed to rationality assumptions; furthering findings against the neutrality of certain anticipated shocks, but still raising questions in regard to the legitimacy of individuals' rational expectations.

A third analysis which probes the Theory of Rational Expectations is Frydman and Rappoport (1987). The researchers review the work of Barro (1977, 1978) and Mishkin (1982a), but implement a vital structural change in their incorporation of anticipated and unanticipated money growth. Frydman and Rappoport install two intercept dummy variables in two different specifications to account for serial fluctuations of the money growth rate.[14] With these two regime shifts integrated into anticipated changes of individuals' expectations over the same time periods, Frydman and Rappoport results suggest that both anticipated and unanticipated money growth affects real output with varying quantitative levels depending on the model. However, none of their models reject any such rationality hypotheses.

C. Longevity Expectations

Hamermesh (1985) was the first paper to examine subjective longevity expectations as predictors of the duration individuals' life spans. Hamermesh employed individual-level data, which was the first of its kind. He randomly selected 650 individuals for polling from the *American Economic Association*'s Biographical Listing of Members from 1978 and sent out questionnaires to 975 individuals randomly chosen from the telephone directory of the medium-sized Midwestern Standard Metropolitan Statistical Areas index. Out of the two data pools, 63%

[14] In one specification the intercept dummy is employed during the third quarter of 1963 in response to systematic and discrete increases in the average money growth rate, and in the other specification the intercept dummy is employed during the third quarter of 1971 in response to the dissolution of the Bretton-Woods System. See Taylor (1983) and Buiter (1983) for a full discussion on these regime changes.

from the *AEA*'s listing and 47% from the SMSA's telephone directory returned usable responses.[15] In examining this dataset, Hamermesh outlines the following qualifications for which to judge the overall distribution of individuals' longevity expectations:

1) Do subjective longevity expectations reflect actuarial survival distributions?

2) Does the researcher deal with demographic variables, specifically age, appropriately? – Does the researcher find older individuals to have higher subjective longevity expectations than younger individuals, ceteris paribus?

3) Do the results conform to the Theory of Rational Expectations? – Are there any systematic biases in their predictions?

4) Do the personal and health regressors make intuitive sense?[16]

In response to his own standards, Hamermesh's results are fairly comparable. In accordance to his first qualification, Hamermesh constructs a Weibull Survival Function[17] to estimate the sample's distribution of individuals' subjective survival probabilities and compares it to an actuarial distribution of males' age 39 survival probabilities.[18] He finds the distribution of individual's subjective survival probabilities to have a higher variance than the actuarial distribution, which is somewhat inconsistent with his first criterion; however, it places a greater importance on the study of longevity expectations as perhaps a more insightful measure of personal health conditions.[19] Further, Hamermesh employs an OLS framework to examine his next three qualifications. He implements individual-level demographic, health, and family

[15] Hamermesh notes that the samples are clearly not representative, but due to the differing features of the two data pools, he states the data will still offer robustness.
[16] For example, would an individual who responded to feeling depressed and lonely with a history of heart problems offer a lower probability of living to age 75 than an individual without those characteristics.
[17] $P_i(t \mid p^s_{60,i}; p^s_{80,i}; x_i) = \exp[-((t - x_i)/\theta_i)^{c_i}], t > x_i$ where $p^s_{60,i}$ is the individual's subjective probability of living to age 60, $p^s_{80,i}$ is the individual's subjective probability of living to age 80, i is a respondent, c_i and θ_i are parameters, and P is the actuarial probability of an individual's survival to age t.
[18] For Men at Age 39, which is the gender and mean age of his sample
[19] Because if the distributions had the same variance, then either type of data (actuarial probabilities or subjective probabilities) could be used for research on personal life-cycle behavior. This finding led Hamermesh to conclude that actuarial probabilities are inappropriate for models of individual level life-cycle decisions.

history variables along with an adjustment variable, DEL_x. This variable is Hamermesh's proxy for testing whether or not individuals' forecasted changes in life expectancy are systematically different from the 1980 life tables from the *Vital Statistics of the United States*. His hypothesis is if DEL_x's coefficient is equal to 1, then rationality holds. Hamermesh finds this specification to satisfy his last three qualifications and support both the legitimate predictive power of longevity expectations and the Theory of Rational Expectations.

Hurd and McGarry (2002) expand on Hamermesh's framework by employing a two wave longitudinal individual-level dataset which included respondents' longevity expectations and reported mortality risks. Hurd and McGarry use a Probit regression model to examine the precursors for death in between waves 1 and 2. They find the correlation between lower longevity expectations and actual death to be statistically significant. After finding evidence for the predictive power of longevity expectations on one's death, Hurd and McGarry then model the determinants of these longevity expectations. Across two specifications: one model with a longevity expectation of living to age 75 and the second with longevity expectations to age 85, the researchers find the following to be significant: age, gender, race, frequency of physical activity, the death of one's parent, the onset of cancer, and one's self-report of their health status in heuristic terms.[20] Although Hurd and McGarry conclude that longevity expectations "show considerable promise for estimating models of decision-making under uncertainty," the researchers fail to account for any amount of attrition within their dataset (Hurd and McGarry, 2002, p. 966). Therefore, their intuitive and theoretically likely findings could potentially be unfounded based upon a biased selection comprised of surviving respondents.[21]

---

[20] Individuals could respond to the following options: very good, good, fair, or poor. Hurd and McGarry then translated these responses into categorical dummy variables.

[21] Hurd and McGarry (2002) do not ignore this selection bias; they do calculate death rates and offer two explanations for their underestimated calculation compared to the 1990-1993 Life Tables: (1) their dataset is only

Smith, Taylor, and Sloan (2001) explicitly disclose their research's close relationship to Hurd and McGarry (2002)'s paper and work to correct their selection bias.[22] To do so, Smith, et al. employ a Probit selection model in accordance with a Heckman (1979) Two-Step Process[23] in waves 1 through 4 of the same dataset to account for survey respondents who either did not respond or if their death could not be verified with certainty. Smith, et al. note that this process solves the dataset's incidental truncation problem evidencing a calculated death rate of 14.50 deaths per thousand as compared to the U.S. Department of Commerce's 1999 Life Table's 13.70 deaths per thousand. After making this correction, Smith et al. run a Probit model to analyze the predictive power of one's longevity expectation in the previous wave on the individual's death in the current wave. They find individuals' longevity expectations, smoking status, health shocks (such as the onset of cancer or lung disease, having a stroke, or contracting diabetes), and lastly, income to be statistically significant at the 5% level. Further, Smith et al. employ the longevity expectation data to test whether or not they exhibit the first half of the Rationality Hypothesis. Smith et al. constructs a Two-Stage Reverse regression model[24] in which they adopt the null hypothesis that all relevant information and factors are fully reflected in the model's longevity prediction from its' first stage. They use Least Squares to fit a regression of an individual's longevity expectation in wave *j*, and then use an individual's death in wave *j+1* as future determinant, hence the term reverse regression. From the results of this test, Smith et al.

---

representative of the non-institutionalized population, which have a lower mortality risk than the entire population, and (2) the time span between waves is not exactly 1, 2, or 3 years, but varies by a couple months in which more deaths could occur.

[22] Smith et al. cite Hurd and McGarry's (1998) unpublished manuscript in their 2001 paper by writing "Michael D. Hurd and Kathleen McGarry (1998) [. . .] are relevant to our analysis" – page 1127

[23] The same process will be employed later in this paper with further explanation.

[24] $\hat{y}_j = a_1 y_j + a_2 d_{j+1} + \varepsilon_j$

reject the null hypothesis and the rationality assumption that individuals' longevity expectations are, on average, fully reflective of all their known and possibly relevant information.

Additionally, two papers, Benitez-Silva and Dwyer (2005) and Benitez-Silva, Dwyer, Gayle, and Muench Benitez-Silva (2008) are related in testing Rational Expectations Theory, but they do not conduct examinations of individuals' longevity expectations, but rather individuals' retirement prospects. The researchers in these two papers examine retirement and educational attainment expectations respectively in much the same methodology of Hamermesh (1985) and Hurd and McGarry (2002), however Benitez-Silva and Dwyer and Benitez-Silva et al. employ a Two Stage Least Squares technique to test rationality. They specify models which analyze the role of new information and health shocks by implementing lagged versions of the forecasted variables as instruments, then control for wellness changes with vectors of health shocks. The researchers conclude that perfect foresight must be rejected in a majority of cases, meaning individuals are not always perfectly accurate. However, they do not find evidence for irrational behavior. In other words, although individuals not may be accurate, they are not systematically biased.

III. Methodology

A. Hypothesis

This paper seeks to answer two important questions regarding individuals' longevity expectations. The first inquiry is, How well-founded are individuals expectations about their life spans? This question is largely based off of Hamermesh (1985)'s original criterions for judging longevity expectations. Therefore, within the context of this hypothesis, we should expect to find our results to align with the positive responses to Hamermesh's qualifications. Namely, individuals’ longevity expectations will first be found to have intuitive reactions to known information. For example, a positive association with weekly exercise, but a negative association with being a current smoker. Secondly, individuals’ longevity expectations should be significant predictors of death.

Provided individuals’ longevity expectations are found to signal death, the second half of our hypothesis will then turn its focus to the potential biases of individuals’ forecasting estimates. This is a de facto test on the Theory of Rational Expectations where the research team assumes a rationality postulate to hold. Therefore, theoretically we should expect to find no systematic biases and that any past information pertinent to the individual will result to be statistically insignificant.

B. Data

All eight existing waves of the Health and Retirement Study (hereon referenced as HRS) were retrieved to provide the basis for this analysis. The HRS is a national panel study of households whose main respondent is within the ages of 45-65. The HRS dataset is collected in biennial waves, with its first wave conducted in 1992. This baseline wave included 12,652 individuals with oversamples of Hispanic-Americans, African-Americans, and Floridians derived from a 0.820 response rate (Juster and Suzman, 1995). These first wave interviews were conducted within individuals’ homes, but subsequent waves’ interviews were completed via telephone. Each in-home interview contained questions covering fourteen sections, including: demographics, physical health, family structure, current work and work history, income, and future plans. Interviews could last up to two hours.

Based upon the extensive nature of the HRS dataset, we were able to collect in-depth information regarding individuals' personal statuses tracked for eight different cross-sections over sixteen years (1992-2006). However, the HRS dataset was not simply employed for its collection of health and personal information, the dataset asks subjective perception questions, which includes a question that produces our variable of interest. We derive of variable of individuals' longevity expectations from individuals' responses to the following HRS question:

> "Using any number from 0 to 100 where 0 equals absolutely no chance and 100 equals absolutely certain, what do you think are the chances you will live to be 75 or more?"[25]

This scaled question bears a strong parallel to Hamermesh's (1985) original longevity study, and is also the variable employed in parts of Smith, et al. (2001) and Hurd and McGarry (2002).

A host of other variables were also retrieved from the HRS. Many of the paper's variables are listed in Table 1.[26]

[25] All questions can found on the HRS's website < http://hrsonline.isr.umich.edu/ >
[26] All data was downloaded via the RAND intercooled STATA file < http://hrsonline.isr.umich.edu/ >

Table 1. Variable Definitions

Variable	Description
$Istat_{i,w}$	Respondent's Interview Status in Wave w
$LE_{i,w}$	Individual's Longevity Expectation in Wave w
$Married_{i,w}$	If individual is married in Wave w then 1; otherwise 0
$Hispanic_{i}$	If individual is Hispanic then 1; otherwise 0
$Black_{i}$	If individual is Black then 1; otherwise 0
$Age_{i,w}$	Is the respondent's age in Wave w
$Male_{i}$	If individual is Male then 1; otherwise 0
$Grad_{i}$	If individual is a college graduate then 1; otherwise 0
$Exercise_{i,w}$	If individual exercise 3+ times a week in Wave w then 1; otherwise 0
$BMI_{i,w}$	Is the individuals BMI calculated from their responded height and weight in Wave w
$Smoke_{i,w}$	If the individual currently smokes in Wave w then 1; otherwise 0
$Earnings_{i,w}$	The sum of the respondent's wage/salary income, bonuses, overtime, commissions, tips, 2nd job, and other forms of income in Wave w
$Hbp_{i,w}$	If the individual has had high blood pressure since the last interview then 1; otherwise 0
$Stroke_{i,w}$	If the individual has had a stroke since the last interview then 1; otherwise 0
$Cancer_{i,w}$	If the individual has had cancer since the last interview then 1; otherwise 0
$Lung_{i,w}$	If the individual has had lung disease since the last interview then 1; otherwise 0
$Psych_{i,w}$	If the individual has had psychological problems since the last interview then 1; otherwise 0
$Heart_{i,w}$	If the individual has had heart disease since the last interview then 1; otherwise 0
$Diab_{i,w}$	If the individual has had diabetes since the last interview then 1; otherwise 0
$Arthri_{i,w}$	If the individual has had arthritis since the last interview then 1; otherwise 0

C. Econometric Problems

The HRS data suffers from two major conditions. The first problem is a product of the longevity question itself. Because the question deliberately sets a horizon – 75 years, the dataset has a different sample than what would be appropriate for the questions of our hypothesis. This non-sampling error is a product of the HRS survey asking its longevity question to all its respondents, instead of just the respondents who could potentially survive to age 75. More suitable procedures could have perhaps been to ask, "To what age do you think you will live?" with the de facto follow-up, "What are the chances that you would survive to that age?" Then track each respondent until they either reached their horizon age or perished. Fortunately, the dataset can be used to obtain a match between the HRS question's horizon age and the actual observed survival condition of the individual. Because a number of the individuals were age 59 or older in 1992 at the start of Wave 1, their longevity expectations can be matched to their actual survival statuses at the age of 75. This allows for the alignment of the forecast horizons with observable and recordable data. This shrinks the number of observations considerably; however, there are still 1,810 respondents where both longevity expectations and the actual survival statuses at age 75 can be obtained. This subsample of the HRS respondents is ultimately employed as the paper's dataset for later testing of Rationality.

The second main problem concerns attrition within the entire dataset and the subsample. Since there is no feasible way of distinguishing all of the respondents who actually died between waves and those that simply did not respond to the survey, one must employ a selection methodology to censor the data appropriately. The two-step logic of Heckman (1979) is employed for this selection bias. A Probit model is used to estimate an inverse Mills ratio, which is included in the corrective methodology's next step as an explanatory variable. White (1982) and Smith et al. (2001) both employ this technique to account for their samples' selection bias. For a comprehensive overview of the Heckman Correction and the Mortality Selection procedure see Appendix 1A.

A minor issue considered is that of Heteroskedasticity. Breusch-Pagan tests on ordinary regressions of the cross-sectional inter-wave data suggest that robust standard errors should be reported. Therefore, all the standard errors shown are robust.

D. Econometric Model

i. Determinants of Longevity Expectations

The first step in this analysis is to answer Hamermesh (1985)'s first and fourth questions and to establish the legitimacy of an individual's longevity expectation in reflecting their actual survival probability. To do so, this paper follows the work of Hamermesh[27], Hurd et al. (2001)[28], and Hurd and McGarry (2002)[29] in analyzing the determinants of one's longevity expectation to substantiate their intuitive bases for later tests of their predictive validity. The model employed to analyze stagnant and historical regressors uses individuals' longevity expectations in the current wave, $LE_{i,w}$, as its dependent variable and is estimated in the following OLS and Tobit forms with a lower bound of 0 and an upper bound of 1:

$$\begin{aligned}LE_{i,w} &= \alpha + \beta_1 Married_{i,w} + \beta_2 Divorced_{i,w} + \beta_3 Widow_{i,w} + \beta_4 Hispanic_i + \beta_5 Black_i + \\ &\beta_6 Age_{i,w} + \beta_7 Male_{i,} + \beta_8 Grad_i + \beta_9 \mathrm{Pr}otes\tan t_{i,w} + \beta_{10} Catholic_{i,w} + \beta_{11} Jewish_{i,w} + \\ &\beta_{12} Exercise_{i,w} + \beta_{13} BMI_{i,w} + \beta_{14} Smoke_{i,w} + \beta_{15} Earnings_{i,w} + \beta_{16} Depressed_{i,w} + \\ &\beta_{17} Happy_{i,w} + \beta_{18} Lonely_{i,w} + \beta_{19} DadAlive_{i,w} + \beta_{20} MomAlive_{i,w} + \beta_{21} \# Kids_{i,w} + \\ &\beta_{22} StressJob_{i,w} + \beta_{23} PhysJob_{i,w} + \beta_{24} Working_{i,w} + \beta_{25} Drink_{i,w} + \beta_{26} South_{i,w} + \varepsilon_i\end{aligned} \quad (2)$$

where the variables included are examined over the longitudinal dataset. Subsequently, to test against more dynamic changes on individuals' longevity expectations, the OLS and Tobit versions of Model (3) is offered with the change in an individual's longevity expectation, $LE_{i,w} - LE_{i,w-1}$, as the dependent variable:

[27] Hamermesh's analysis focuses on age. He includes dummy variables for the ages of the respondents' parents and grandparents along with his controls of individuals' personal responses.
[28] Hurd et al. employ a different data, the Asset and Health Dynamics Study (AHEAD), to examine longevity expectations amongst the elderly
[29] Hurd and McGarry (2002) use a similar approach to Hurd et al. with HRS data by including the death of one's parents, health shock information, wealth, education, and other demographic variables.

$$LE_{i,w} - LE_{i,w-1} = \alpha + \beta_1 Married_{i,w} + \beta_2 Divorced_{i,w} + \beta_3 Widow_{i,w} + \beta_4 Hispanic_i +$$
$$\beta_5 Black_i + \beta_6 Age_{i,w} + \beta_7 Male_{i,} + \beta_8 Grad_i + \beta_9 \mathrm{Pr}otes\tan t_{i,w} + \beta_{10} Catholic_{i,w} +$$
$$\beta_{11} Jewish_{i,w} + \beta_{12} Exercise_{i,w} + \beta_{13} BMI_{i,w} + \beta_{14} Smoke_{i,w} + \beta_{15} Earnings_{i,w} +$$
$$\beta_{16} Depressed_{i,w} + \beta_{17} Happy_{i,w} + \beta_{18} Lonely_{i,w} + \beta_{19} DadAlive_{i,w} + \beta_{20} MomAlive_{i,w} + \quad (3)$$
$$\beta_{21} \# Kids_{i,w} + \beta_{22} StressJob_{i,w} + \beta_{23} PhysJob_{i,w} + \beta_{24} Working_{i,w} + \beta_{25} Drink_{i,w} +$$
$$\beta_{26} South_{i,w} + \beta_{27} Hbp_{i,w} \beta_{28} Stroke_{i,w} + \beta_{29} Cancer_{i,w} + \beta_{30} Lung_{i,w} \beta_{31} Psych_{i,w} +$$
$$\beta_{32} Heart_{i,w} + \beta_{33} Diab_{i,w} + \beta_{34} Arthri_{i,w} + \varepsilon_i$$

The variables of Model (2) and (3) are described in Table 2.

Table 2. Variable Definitions

Variable	Description
$LE_{i,w}$	Individuals' Longevity Expectation in Wave w
$LE_{i,w-1}$	Individual’s Longevity Expectation in the prior Wave
$Age_{i,w}$	Is the respondent's age in Wave w
$Hispanic_{i}$	If individual is Hispanic then 1; otherwise 0
$Black_{i}$	If individual is Black then 1; otherwise 0
$Male_{i}$	If individual is Male then 1; otherwise 0
$Married_{i,w}$	If individual is married in Wave w then 1; otherwise 0
$Grad_{i}$	If individual is a college graduate then 1; otherwise 0
$Exercise_{i,w}$	If individual exercise 3+ times a week in Wave w then 1; otherwise 0
$BMI_{i,w}$	Is the individuals BMI calculated from their responded height and weight in Wave w
$Smoke_{i,w}$	If the individual currently smokes in Wave w then 1; otherwise 0
$Earnings_{i,w}$	The sum of the respondent's wage/salary income, bonuses, overtime, commissions, tips, 2nd job, and other forms of income in Wave w
$Depressed_{i,w}$	If individual has felt depressed leading up to Wave w’s interview then 1; otherwise 0
$Happy_{i,w}$	If individual has felt happy leading up to Wave w’s interview then 1; otherwise 0
$Lonely_{i,w}$	If individual has felt lonely leading up to Wave w’s interview then 1; otherwise 0
$DadAlive_{i,w}$	If individual’s father is alive in Wave w then 1; otherwise 0
$MomAlive_{i,w}$	If individual’s mother is alive in Wave w then 1; otherwise 0
$Working_{i,w}$	If individual holds some type of paid/compensated job in Wave w then 1; otherwise 0
$\#Kids_{i,w}$	The number of children the respondent lists as living in Wave w
$Smoke_{i,w}$	If the individual currently smokes in Wave w then 1; otherwise 0
$Hbp_{i,w}$	If the individual has had high blood pressure since the last interview then 1; otherwise 0
$Stroke_{i,w}$	If the individual has had a stroke since the last interview then 1; otherwise 0
$Cancer_{i,w}$	If the individual has had cancer since the last interview then 1; otherwise 0
$Lung_{i,w}$	If the individual has had lung disease since the last interview then 1; otherwise 0
$Psych_{i,w}$	If the individual has had psychological problems since the last interview then 1; otherwise 0
$Heart_{i,w}$	If the individual has had heart disease since the last interview then 1; otherwise 0
$Diab_{i,w}$	If the individual has had diabetes since the last interview then 1; otherwise 0
$Arthri_{i,w}$	If the individual has had arthritis since the last interview then 1; otherwise 0

ii. Longevity Expectations' relationship with Death

After establishing predictive validity of one's longevity expectation based off of their known personal, demographic, and health information, this paper then seeks to analyze the relationship between individuals' longevity expectations and deaths. This test furthers the application of implementing individuals' longevity expectations as it extends the variable's predictive power from a personal health context to larger life cycle models, such as in Friedman (1957) and Modigliani (1963). A simple Probit model derived from Smith et al. (2001)'s framework is employed to test the ability of individuals to signal their own deaths through their subjective longevity expectations. Model (4) is offered for these ends:

$$Alive_{i,w} = \alpha + \beta_1 LE_{i,w-1} + \beta_2 Smoke_{i,w-1} + \beta_3 Drink_{i,w-1} + \beta_4 Earnings_{i,w-1} + \beta_5 \Omega_{i,w-1} + \varepsilon_i \quad (4)$$

where $Alive_{i,w}$ is a 1 if the individual is listed as alive in the current wave and 0 otherwise, $LE_{i,w-1}$ is an individual's longevity expectation from one wave prior, $Smoke_{i,w-1}$ is the individual's smoking status in the prior wave; 1 if they do and 0 otherwise, $Drink_{i,w-1}$ is the individual's alcoholic drinking status in the prior wave; 1 if they do and 0 otherwise, and $Earnings_{i,w-1}$ is the individual's total earnings from the prior wave. Lastly, health shocks to an individual are still estimated separately; however, the notation, $\Omega_{i,w-1}$, denotes a vector of the eight conditions from one wave prior that were employed in previous models.

iii. Testing Aggregate Rationality

Recalling Muth (1961)'s original Theory of Rational Expectations from equation (1), a model which examines a rationality assumption must test the relationship between anticipated information and the expectation, as well as the relationship between news or shocks and the expectation. To do so, the framework of Davies and Lahiri (1995) and (1999) is combined with that of Muth's. Let $Alive75_{i,T}$ be the actual survival status of the ith individual at the individual's target life duration of 75 years. Each individual's actual survival status at target time 75 years is then dependent on:

1.) Known or possibly anticipated information in Wave *w*,

2.) News or shocks – unanticipated information that occurs between Wave, *w*, and the horizon target time, *T*,

3) A normally distributed random component, ε,

which yields individual i's actual survival status at target time, *T*, to be:

$$\text{Alive75}_{\text{i,T}} = \beta_1 \Omega_{i,w} + \beta_2 \lambda_{i,T-w} + \varepsilon_i \tag{5}$$

where **Ω** is known and observable information in wave *w*, $\lambda_{i,T\text{-}w}$ is news and shocks between wave *w* and the target time, and ε is the error term.

Further, an individual's expectation of their survival status at age 75 follows much the same framework, except individuals can only forecast their longevity expectation at a time in wave *w* based upon known information, **Ω**, and preexisting biases, φ, which could be thought of as innate optimism or pessimism:

$$LE_{i,w} = \beta_1 \Omega_{i,w} - \varphi_{i,w} + \upsilon_i \tag{6}$$

Subtracting (5) from (6) yields an expression for forecast errors where longevity expectations, which assumedly already have incorporated known information, differ from actual survival status due to preexisting biases, news and shocks, observable information, and stochastic errors:

$$Alive75_{i,w} - LE_{i,w} = \varphi_{i,w} + \beta_2 \lambda_{i,T-w} + \beta_3 \Omega_{i,w} + \mu_i \tag{7}$$

with μ_i serving as a composite error term. Thus a minimally-specified test for rationality holds when the bias term, $\varphi_{i,w}$, is not statistically different from zero, and a more comprehensive form holds when $\varphi_{i,w}$ and β_3 are jointly not statistically different from zero.

IV. Results

A.) Determinants of Longevity Expectations

The OLS and Tobit regression results and each variable's expected association from Model (2) can be found in Table 3 on the next two pages.

Table 3. The Effect of Mortality Risks on Longevity Expectations (Waves 1-4)

	LE in Wave 1		LE in Wave 2		LE in Wave 3		LE in Wave 4		Exp. Sign
Variable	OLS	Tobit	OLS	Tobit	OLS	Tobit	OLS	Tobit	
$LE_{i,w-1}$			0.471***	0.593***	0.509***	0.677***	0.472***	0.593***	+
			(0.010)	(0.014)	(0.012)	(0.016)	(0.013)	(0.016)	
$Hbp_{i,w}$	-0.033***	-0.044***	-0.002	-0.007	0.006	0.013	-0.008	-0.016	-
	(0.006)	(0.008)	(0.015)	(0.020)	(0.018)	(0.025)	(0.015)	(0.019)	
$Stroke_{i,w}$	-0.034*	-0.042*	-0.073*	-0.087	-0.112**	-0.152**	-0.032	-0.048	-
	(0.018)	(0.024)	(0.043)	(0.057)	(0.047)	(0.064)	(0.047)	(0.058)	
$Cancer_{i,w}$	-0.066***	-0.082***	-0.119***	-0.149***	-0.068**	-0.089**	-0.098***	-0.112***	-
	(0.012)	(0.016)	(0.028)	(0.035)	(0.031)	(0.042)	(0.029)	(0.037)	
$Lung_{i,w}$	-0.078***	-0.099***	-0.078***	-0.097***	-0.073*	-0.098*	-0.063**	-0.081**	-
	(0.011)	(0.014)	(0.027)	(0.035)	(0.041)	(0.057)	(0.030)	(0.039)	
$Psych_{i,w}$	-0.023**	-0.033**	0.025	0.026	-0.064**	-0.090*	-0.035	-0.046	-
	(0.009)	(0.013)	(0.019)	(0.025)	(0.028)	(0.038)	(0.025)	(0.032)	
$Heart_{i,w}$	-0.077***	-0.103***	-0.048***	-0.055**	-0.077*	-0.110***	-0.026	-0.023	-
	(0.008)	(0.011)	(0.018)	(0.023)	(0.022)	(0.028)	(0.021)	(0.028)	
$Diab_{i,w}$	-0.049***	-0.063***	-0.048**	-0.068**	-0.007	-0.017	-0.002	0.000	-
	(0.009)	(0.013)	(0.023)	(0.030)	(0.024)	(0.034)	(0.026)	(0.033)	
$Arthri_{i,w}$	-0.024***	-0.032***	-0.019	-0.026*	0.007	0.005	0.022	0.036*	-
	(0.006)	(0.008)	(0.012)	(0.015)	(0.013)	(0.018)	(0.014)	(0.019)	
$Married_{i,w}$	0.007	0.006	-0.005	-0.012	0.007	0.001	0.005	0.006	+
	(0.012)	(0.017)	(0.012)	(0.016)	(0.014)	(0.020)	(0.015)	(0.020)	
$Divorce_{i,w}$	0.012	0.015	0.000	-0.002	-0.001	-0.010	0.018	0.024	-
	(0.014)	(0.019)	(0.014)	(0.019)	(0.016)	(0.023)	(0.017)	(0.023)	
$Widow_{i,w}$	0.001	-0.003	-0.004	-0.003	-0.001	-0.009	0.027	0.041	-
	(0.016)	(0.023)	(0.016)	(0.022)	(0.018)	(0.025)	(0.019)	(0.025)	
$Hispanic_{i}$	-0.073***	-0.096***	-0.021*	-0.025*	-0.047***	-0.064***	-0.047***	-0.055***	-
	(0.011)	(0.016)	(0.012)	(0.015)	(0.013)	(0.018)	(0.015)	(0.020)	
$Black_{i}$	0.039***	0.061***	0.0293***	0.050***	0.005	0.024	0.023**	0.043***	-
	(0.008)	(0.011)	(0.009)	(0.012)	(0.010)	(0.015)	(0.011)	(0.014)	
$Age_{i,w}$	0.006***	0.008***	0.002***	0.003***	0.001**	0.003***	0.002**	0.002**	+
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
$Male_{i}$	-0.056***	-0.071***	-0.016***	-0.018**	-0.031***	-0.036***	-0.030***	-0.035***	-
	(0.006)	(0.008)	(0.006)	(0.008)	(0.007)	(0.009)	(0.007)	(0.009)	
$Grad_{i}$	0.033***	0.024***	0.0169***	0.007	0.026***	0.011	0.020***	0.000	+
	(0.006)	(0.009)	(0.006)	(0.008)	(0.007)	(0.009)	(0.007)	(0.009)	
$Protestant_{i,w}$	0.017	0.022	0.006	0.009	0.003	-0.003	0.015	0.014	+
	(0.012)	(0.016)	(0.011)	(0.015)	(0.013)	(0.018)	(0.013)	(0.017)	
$Catholic_{i,w}$	0.009	0.013	0.020*	0.028*	0.001	-0.001	0.024*	0.032*	+
	(0.012)	(0.017)	(0.012)	(0.015)	(0.013)	(0.019)	(0.014)	(0.018)	
$Jewish_{i,w}$	0.018	0.028	0.039*	0.054**	-0.016	-0.026	0.017	0.027	+
	(0.021)	(0.029)	(0.020)	(0.027)	(0.025)	(0.033)	(0.023)	(0.031)	
$Exercise_{i,w}$	0.031***	0.046***	0.029***	0.037***	0.034***	0.045***	0.027***	0.036***	+
	(0.006)	(0.009)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	
$BMI_{i,w}$	-0.001	-0.001	-0.001***	-0.002**	-0.001**	-0.001*	-0.002***	-0.003***	-
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
$Smoke_{i,w}$	-0.047***	-0.057***	-0.031***	-0.034***	-0.046***	-0.058***	-0.039***	-0.041***	-
	(0.006)	(0.009)	(0.007)	(0.009)	(0.008)	(0.010)	(0.008)	(0.011)	
$Earnings_{i,w}$	0.000	0.000	0.002***	0.003**	0.000	0.000	0.000*	0.000	+
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	
$Depressed_{i,w}$	-0.051***	-0.067***	0.054*	-0.021	-0.044	-0.055***	-0.046***	-0.052***	-
	(0.007)	(0.009)	(0.010)	(0.014)	(0.012)	(0.016)	(0.011)	(0.014)	
$Happy_{i,w}$	0.089***	0.116***	0.054***	0.071***	0.047	0.064***	0.011	0.000	+
	(0.017)	(0.024)	(0.011)	(0.014)	(0.012)	(0.017)	(0.012)	(0.015)	
$Lonely_{i,w}$	-0.024***	-0.032***	-0.028***	-0.037***	-0.018	-0.025	-0.027**	-0.036**	-
	(0.007)	(0.011)	(0.011)	(0.014)	(0.011)	(0.015)	(0.011)	(0.014)	
$DadAlive_{i,w}$	0.039***	0.052***	0.019***	0.020**	0.017**	0.021*	0.025***	0.028**	+
	(0.007)	(0.009)	(0.007)	(0.009)	(0.008)	(0.012)	(0.009)	(0.012)	
$MomAlive_{i,w}$	0.054***	0.067***	0.018***	0.019***	0.023***	0.023***	0.025***	0.027***	+
	(0.005)	(0.007)	(0.005)	(0.007)	(0.006)	(0.009)	(0.006)	(0.008)	
$Kids_{i,w}$	0.002*	0.003*	-0.003**	-0.003*	0.000	0.001	0.001	0.001	+
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	
$StressJob_{i,w}$	-0.021***	-0.031***	-0.004	-0.01	-0.010	-0.013	0.003	0.003	-
	(0.006)	(0.008)	(0.006)	(0.008)	(0.007)	(0.010)	(0.008)	(0.010)	
$PhysJob_{i,w}$	-0.019***	-0.022**	0.004	0.011	-0.008	-0.004	-0.001	0.001	-
	(0.007)	(0.009)	(0.007)	(0.008)	(0.007)	(0.010)	(0.008)	(0.010)	
$Working_{i,w}$	0.049***	0.064***	0.014*	0.017	0.039***	0.047***	0.009	0.008	+
	(0.008)	(0.012)	(0.008)	(0.011)	(0.009)	(0.012)	(0.009)	(0.012)	
$Drink_{i,w}$	0.036***	0.046***	0.025***	0.029***	0.018***	0.018**	0.018***	0.016*	+
	(0.006)	(0.008)	(0.006)	(0.007)	(0.006)	(0.008)	(0.006)	(0.008)	
$South_{i,w}$	-0.011*	-0.013*	-0.009*	-0.010	-0.019***	-0.020**	-0.010	0.000	+
	(0.005)	(0.008)	(0.005)	(0.007)	(0.006)	(0.008)	(0.006)	(0.008)	
R-squared	0.278	0.086	0.242	0.282	0.249	0.265	0.236	0.312	
Observations	11557	11557	9055	9055	7511	7511	6245	6245	

Dependent variable is individual-reported expected probability of living to the age of 75 in each wave. Coefficients with robust standard errors in parenthesis reported. For Tobit regressions, Pseudo R-squared is reported. Constant not reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Table 3 (continued). The Effect of Mortality Risks on Longevity Expectations (Waves 5-8)

	LE in Wave 5		LE in Wave 6		LE in Wave 7		LE in Wave 8		Exp. Sign
Variable	OLS	Tobit	OLS	Tobit	OLS	Tobit	OLS	Tobit	
$LE_{i,w-1}$	0.508***	0.621***	0.533***	0.655***	0.531***	0.641***	0.517***	0.616***	+
	(0.013)	(0.017)	(0.016)	(0.021)	(0.018)	(0.022)	(0.027)	(0.033)	
$Hbp_{i,w}$	0.007	0.009	0.007	0.008	-0.031*	-0.039**	-0.025	-0.027	-
	(0.015)	(0.019)	(0.016)	(0.020)	(0.016)	(0.020)	(0.030)	(0.035)	
$Stroke_{i,w}$	-0.110***	-0.120**	-0.056*	-0.092**	-0.058	-0.072	-0.026	-0.033	-
	(0.037)	(0.048)	(0.034)	(0.039)	(0.041)	(0.052)	(0.069)	(0.083)	
$Cancer_{i,w}$	-0.044**	-0.051*	-0.055*	-0.065*	-0.083***	-0.106***	-0.100	-0.119	-
	(0.022)	(0.028)	(0.032)	(0.038)	(0.031)	(0.038)	(0.066)	(0.079)	
$Lung_{i,w}$	-0.103***	-0.128***	0.011	0.017	-0.063*	-0.076*	-0.026	-0.023	-
	(0.031)	(0.041)	(0.031)	(0.040)	(0.032)	(0.039)	(0.039)	(0.050)	
$Psych_{i,w}$	-0.012	-0.015	-0.035	-0.047	0.006	0.019	-0.025	-0.022	-
	(0.026)	(0.031)	(0.027)	(0.032)	(0.029)	(0.037)	(0.044)	(0.054)	
$Heart_{i,w}$	-0.030	-0.023	-0.020	-0.006	-0.040	-0.056*	-0.030	-0.045	-
	(0.021)	(0.027)	(0.026)	(0.033)	(0.026)	(0.031)	(0.044)	(0.053)	
$Diab_{i,w}$	-0.011	-0.010	0.009	0.011	0.013	0.021	-0.107***	-0.117**	-
	(0.024)	(0.029)	(0.022)	(0.027)	(0.028)	(0.036)	(0.041)	(0.046)	
$Arthri_{i,w}$	-0.029*	-0.039*	-0.03**	-0.042**	-0.006	-0.011	-0.022	-0.021	-
	(0.017)	(0.020)	(0.015)	(0.019)	(0.017)	(0.022)	(0.030)	(0.035)	
$Married_{i,w}$	0.012	0.014	0.000	-0.002	0.002	-0.002	0.022	0.018	+
	(0.015)	(0.019)	(0.017)	(0.021)	(0.018)	(0.023)	(0.029)	(0.034)	
$Divorce_{i,w}$	0.004	0.012	0.007	0.009	-0.005	-0.009	0.018	0.012	-
	(0.017)	(0.022)	(0.020)	(0.025)	(0.022)	(0.027)	(0.036)	(0.042)	
$Widow_{i,w}$	-0.005	-0.005	0.007	0.013	-0.014	-0.019	-0.004	-0.017	-
	(0.018)	(0.023)	(0.022)	(0.027)	(0.023)	(0.029)	(0.036)	(0.043)	
$Hispanic_{i}$	-0.067	-0.077***	-0.004	-0.003	-0.05***	-0.064***	-0.094***	-0.108***	-
	(0.016)	(0.020)	(0.017)	(0.022)	(0.018)	(0.022)	(0.032)	(0.039)	
$Black_{i}$	0.021	0.038***	0.008	0.028*	0.033**	0.059***	0.039*	0.061**	-
	(0.011)	(0.014)	(0.013)	(0.017)	(0.014)	(0.018)	(0.021)	(0.027)	
$Age_{i,w}$	0.002***	0.003***	0.003***	0.004***	0.002**	0.003**	0.004**	0.004**	+
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	
$Male_{i}$	-0.015**	-0.015*	-0.025***	-0.028***	-0.034***	-0.036***	-0.022	-0.021	-
	(0.007)	(0.009)	(0.008)	(0.010)	(0.009)	(0.012)	(0.017)	(0.020)	
$Grad_{i}$	0.011	-0.001	0.027***	0.018*	-0.012	-0.028**	0.025*	0.013	+
	(0.007)	(0.009)	(0.009)	(0.011)	(0.010)	(0.012)	(0.015)	(0.018)	
$Protestant_{i,w}$	-0.014	-0.023	0.000	-0.005	-0.023	-0.036*	-0.023	-0.036	+
	(0.013)	(0.017)	(0.015)	(0.019)	(0.016)	(0.021)	(0.024)	(0.030)	
$Catholic_{i,w}$	-0.008	-0.014	-0.007	-0.008	-0.022	-0.026	-0.008	-0.019	+
	(0.014)	(0.018)	(0.016)	(0.020)	(0.018)	(0.022)	(0.027)	(0.033)	
$Jewish_{i,w}$	-0.025	-0.026	0.007	0.019	0.023	0.035	0.005	0.002	+
	(0.026)	(0.033)	(0.031)	(0.039)	(0.033)	(0.043)	(0.049)	(0.062)	
$Exercise_{i,w}$	0.019***	0.023***	0.024**	0.031***	0.018**	0.019*	0.041***	0.042***	+
	(0.006)	(0.008)	(0.007)	(0.009)	(0.009)	(0.011)	(0.013)	(0.016)	
$BMI_{i,w}$	-0.002***	-0.002**	0.000	0.000	-0.002**	-0.001	0.000	0.000	-
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	
$Smoke_{i,w}$	-0.031***	-0.028**	-0.035***	-0.035***	-0.046***	-0.051***	0.010	0.012	-
	(0.009)	(0.011)	(0.010)	(0.013)	(0.012)	(0.015)	(0.017)	(0.021)	
$Earnings_{i,w}$	0.000	0.000	-0.001	0.000	0.006***	0.000	0.000	0.000	+
	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)	
$Depressed_{i,w}$	-0.030**	-0.036**	-0.032**	-0.038**	-0.047***	-0.050**	-0.029	-0.036	-
	(0.012)	(0.015)	(0.015)	(0.018)	(0.016)	(0.019)	(0.026)	(0.031)	
$Happy_{i,w}$	0.023*	0.029*	0.019	0.025	0.036**	0.048**	0.077***	0.093***	+
	(0.013)	(0.016)	(0.013)	(0.016)	(0.016)	(0.020)	(0.025)	(0.029)	
$Lonely_{i,w}$	-0.008	-0.009	-0.046***	-0.053***	-0.019	-0.023	0.030	0.037	-
	(0.011)	(0.014)	(0.013)	(0.017)	(0.016)	(0.019)	(0.026)	(0.032)	
$DadAlive_{i,w}$	0.014	0.013	0.024**	0.026**	0.010	0.014	0.004	0.001	+
	(0.010)	(0.013)	(0.012)	(0.015)	(0.013)	(0.017)	(0.020)	(0.025)	
$MomAlive_{i,w}$	0.021***	0.024***	0.020**	0.019**	0.020**	0.022*	0.014	0.015	+
	(0.007)	(0.009)	(0.008)	(0.010)	(0.009)	(0.011)	(0.015)	(0.018)	
$Kids_{i,w}$	-0.001	0.000	0.000	0.000	-0.002	-0.003	-0.001	-0.001	+
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	
$StressJob_{i,w}$	-0.002	-0.003	0.008	0.016	-0.022*	-0.026**	-0.001	-0.005	-
	(0.008)	(0.010)	(0.010)	(0.012)	(0.011)	(0.014)	(0.018)	(0.021)	
$PhysJob_{i,w}$	-0.006	-0.005	-0.001	0.003	0.002	0.009	0.010	0.013	-
	(0.008)	(0.010)	(0.010)	(0.013)	(0.011)	(0.014)	(0.018)	(0.021)	
$Working_{i,w}$	0.027***	0.032***	0.018*	0.017	-0.001	-0.001	0.015	0.014	+
	(0.009)	(0.012)	(0.011)	(0.014)	(0.012)	(0.015)	(0.020)	(0.024)	
$Drink_{i,w}$	0.002	-0.002	0.020***	0.022**	0.001	-0.005	0.027**	0.030*	+
	(0.007)	(0.008)	(0.008)	(0.010)	(0.008)	(0.010)	(0.013)	(0.016)	
$South_{i,w}$	-0.002	0.000	-0.014*	-0.014	0.015*	0.024**	0.011	0.016	+
	(0.007)	(0.008)	(0.008)	(0.009)	(0.008)	(0.010)	(0.013)	(0.016)	
R-squared	0.224	0.379	0.223	0.407	0.219	0.433	0.232	0.407	
Observations	5265	5265	3995	3995	3039	0.039	1400	1400	

Dependent variable is individual-reported expected probability of living to the age of 75 in each wave. Coefficients with robust standard errors in parenthesis reported. For Tobit regressions, Pseudo R-squared is reported. Constant not reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

In noticing the column titled ‘Exp. Sign’, the results seem to largely indicate that individuals’ longevity expectations are reflective of the actual mortality risks facing individuals. The only variable that is consistently counter-intuitive is the dummy variable, $Black_i$, demarcating 1 if an individual is African-American and 0 otherwise. The positive association of $Black_i$ can be interpreted as African-Americans on average expect to live longer than individuals who are not African-American. A possible reason for this optimism is that the HRS dataset does not survey institutionalized individuals, therefore African-Americans who do make it into the sample use their status as non-institutionalized as further information to add to their longevity expectation.

Other intriguing results found from the estimation are the changes in association through different waves of variables, $Kids_{i,w}$ and $South_{i,w}$, and the insignificance of marital status and religion on individuals’ longevity expectations. Nevertheless, the first research question is principally answered, and Hamermesh (1985)’s first, second, and fourth criterions are satisfied; individuals’ longevity expectations are reflective of actual mortality risks.

The estimation of the more dynamic features of individuals’ longevity expectations through Model (3) can be found in Table 4 on the next two pages.

Table 4. The Effect of Mortality Risks on Changes in Longevity Expectations (Waves 1-5)

	LE2 - LE1		LE3 - LE2		LE4 - LE3		LE5 - LE4		Exp. Sign
Variable	OLS	Tobit	OLS	Tobit	OLS	Tobit	OLS	Tobit	
$Hbp_{i,w}$	0.011	0.001	-0.004	-0.002	-0.011	-0.011	0.009	0.016	-
	(0.017)	(0.028)	(0.021)	(0.031)	(0.018)	(0.032)	(0.017)	(0.029)	
$Stroke_{i,w}$	-0.065	-0.007	-0.057	-0.051	-0.015	0.019	-0.087**	-0.089	-
	(0.051)	(0.068)	(0.054)	(0.083)	(0.062)	(0.103)	(0.036)	(0.061)	
$Cancer_{i,w}$	-0.107***	-0.095**	-0.062*	-0.030	-0.099***	-0.066	-0.057**	-0.101**	-
	(0.032)	(0.047)	(0.034)	(0.047)	(0.034)	(0.054)	(0.025)	(0.047)	
$Lung_{i,w}$	-0.056*	0.000	-0.083**	-0.013	-0.022**	-0.074	-0.062*	-0.091	-
	(0.033)	(0.045)	(0.043)	(0.068)	(0.035)	(0.066)	(0.033)	(0.056)	
$Psych_{i,w}$	0.028	0.027	-0.052*	0.069	-0.024	-0.031	-0.005	-0.003	-
	(0.022)	(0.034)	(0.031)	(0.047)	(0.029)	(0.051)	(0.029)	(0.048)	
$Heart_{i,w}$	-0.033	-0.037	-0.055**	-0.086	-0.021**	-0.039	-0.035	-0.031	-
	(0.022)	(0.035)	(0.024)	(0.038)	(0.023)	(0.045)	(0.023)	(0.037)	
$Diab_{i,w}$	-0.046*	-0.019	-0.015	-0.037	0.021	0.032	0.024	0.061	-
	(0.027)	(0.039)	(0.026)	(0.043)	(0.028)	(0.051)	(0.027)	(0.041)	
$Arthri_{i,w}$	-0.017	-0.015	0.006	0.022	0.012	0.006	-0.038*	-0.037	-
	(0.014)	(0.022)	(0.015)	(0.022)	(0.015)	(0.028)	(0.019)	(0.032)	
$Married_{i,w}$	-0.002	0.005	0.012	0.017	0.007	-0.017	0.002	0.003	+
	(0.015)	(0.023)	(0.016)	(0.024)	(0.018)	(0.029)	(0.016)	(0.028)	
$Divorce_{i,w}$	0.002	0.015	-0.003	0.005	0.026	0.029	-0.001	-0.003	-
	(0.017)	(0.027)	(0.018)	(0.024)	(0.021)	(0.033)	(0.019)	(0.032)	
$Widow_{i,w}$	0.011	0.042	0.006	-0.004	0.039*	0.027	-0.021	-0.042	-
	(0.019)	(0.031)	(0.021)	(0.031)	(0.022)	(0.036)	(0.021)	(0.035)	
$Hispanic_{i}$	0.016	0.038*	-0.014	0.025	-0.005	0.066**	-0.026	0.021	0
	(0.014)	(0.021)	(0.015)	(0.022)	(0.018)	(0.027)	(0.018)	(0.028)	
$Black_{i}$	0.009	0.039**	-0.017	0.003	0.004	0.026	0.002	0.011	0
	(0.011)	(0.015)	(0.012)	(0.016)	(0.012)	(0.021)	(0.012)	(0.021)	
$Age_{i,w}$	0.000	0.000	0.000	-0.001	0.000	-0.001	0.000	0.001	0
	0.000	0.000	0.000	(0.001)	0.000	(0.001)	0.000	(0.002)	
$Male_{i}$	0.015**	0.032***	-0.011	-0.011	0.000	0.013	0.011	0.018	0
	(0.006)	(0.011)	(0.007)	(0.011)	(0.007)	(0.014)	(0.008)	(0.014)	
$Grad_{i}$	0.000	-0.015	0.008	0.003	0.000	-0.021	-0.008	-0.021	0
	(0.007)	(0.013)	(0.007)	(0.012)	(0.008)	(0.016)	(0.008)	(0.015)	
$Protestant_{i,w}$	-0.007	0.003	-0.003	-0.004	0.009	0.051*	-0.015	-0.002	+
	(0.012)	(0.022)	(0.014)	(0.022)	(0.014)	(0.028)	(0.015)	(0.026)	
$Catholic_{i,w}$	0.011	0.034	-0.014	0.019	0.014	0.053*	-0.014	-0.003	+
	(0.013)	(0.023)	(0.015)	(0.024)	(0.015)	(0.031)	(0.015)	(0.028)	
$Jewish_{i,w}$	0.021	0.055	-0.043	-0.067	0.007	0.049	-0.032	-0.023	+
	(0.023)	(0.041)	(0.027)	(0.043)	(0.024)	(0.051)	(0.029)	(0.052)	
$Exercise_{i,w}$	0.007	-0.001	0.007	0.002	0.005	0.001	-0.009	-0.027**	+
	(0.007)	(0.012)	(0.007)	(0.011)	(0.007)	(0.013)	(0.007)	(0.013)	
$BMI_{i,w}$	0.000	0.000	0.000	0.001	-0.001*	-0.001	-0.001	-0.001	-
	0.000	0.000	0.000	(0.001)	0.000	(0.001)	(0.001)	(0.001)	
$Smoke_{i,w}$	-0.005	0.012	0.022**	-0.023	-0.007	0.012	0.005	0.002	-
	(0.008)	(0.012)	(0.009)	(0.013)	(0.009)	(0.017)	(0.009)	(0.016)	
$Earnings_{i,w}$	0.001	0.000	0.000	-0.002	0.001**	0.001**	0.000	0.002	+
	(0.001)	(0.001)	0.000	(0.001)	0.000	0.000	0.000	(0.002)	
$Depressed_{i,w}$	0.003	0.051***	-0.017	-0.005	-0.005	0.014	0.003	0.002	-
	(0.012)	(0.018)	(0.013)	(0.019)	(0.013)	(0.021)	(0.013)	(0.022)	
$Happy_{i,w}$	0.014	0.061***	0.031**	0.035	-0.011*	-0.018	0.009	0.039	+
	(0.013)	(0.021)	(0.014)	(0.021)	(0.013)	(0.022)	(0.014)	(0.024)	
$Lonely_{i,w}$	-0.025**	-0.019	-0.005	0.005	-0.018	0.001	0.007	0.038*	-
	(0.013)	(0.019)	(0.013)	(0.018)	(0.013)	(0.021)	(0.013)	(0.021)	
$DadAlive_{i,w}$	-0.001	-0.002	-0.012	-0.013	0.004	-0.001	-0.007	-0.005	+
	(0.008)	(0.014)	(0.009)	(0.015)	(0.011)	(0.019)	(0.012)	(0.021)	
$MomAlive_{i,w}$	-0.009	-0.028***	0.003	-0.001	-0.001	-0.007	-0.002	-0.009	+
	(0.006)	(0.011)	(0.007)	(0.011)	(0.007)	(0.014)	(0.008)	(0.013)	
$Kids_{i,w}$	-0.003**	-0.002	-0.001	-0.002	0.001	0.005	-0.001	0.001	+
	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)	
$StressJob_{i,w}$	0.002	0.003	-0.004	-0.012	0.009	0.035**	-0.005	-0.019	-
	(0.007)	(0.013)	(0.008)	(0.012)	(0.009)	(0.017)	(0.009)	(0.016)	
$PhysJob_{i,w}$	0.004	0.007	-0.002	-0.001	0.001	0.015	-0.004	0.022	-
	(0.007)	(0.013)	(0.008)	(0.013)	(0.008)	(0.017)	(0.009)	(0.017)	
$Working_{i,w}$	-0.005	-0.017	0.014	0.006	-0.018*	-0.058***	0.012	-0.014	+
	(0.009)	(0.016)	(0.011)	(0.016)	(0.010)	(0.020)	(0.011)	(0.019)	
$Drink_{i,w}$	0.000	-0.017	0.002	0.001	-0.013*	-0.028**	-0.019***	-0.024**	+
	(0.006)	(0.011)	(0.007)	(0.011)	(0.007)	(0.013)	(0.007)	(0.013)	
$South_{i,w}$	-0.002	-0.001	-0.011	-0.001	0.003	0.023*	0.008	0.032**	+
	(0.006)	(0.011)	(0.007)	(0.011)	(0.007)	(0.013)	(0.007)	(0.013)	
R-squared	0.283	0.008	0.283	0.005	0.279	0.009	0.261	0.009	
Observations	9055	9055	7511	7511	6245	6245	5265	5265	

Dependent variable is the change in the individual-reported expected probability of living to the age of 75 from wave to wave. For example "LE2 - LE1" represents the reported probability in wave 2 minus the reported probability in wave 1. Coefficients with robust standard errors in parenthesis reported. For Tobit regressions, Pseudo R-squared is reported. Constant not reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Table 4 (contin). The Effect of Mortality Risks on Longevity Expectations (Waves 5-8)

	LE6 - LE5		LE7 - LE6		LE8 - LE7		Exp. Sign
Variable	OLS	Tobit	OLS	Tobit	OLS	Tobit	
Hbp $_{i,w}$	0.008 (0.018)	-0.001 (0.032)	-0.024 (0.017)	-0.057* (0.033)	-0.021 (0.033)	0.005 (0.052)	-
Stroke $_{i,w}$	-0.051 (0.039)	-0.082 (0.069)	-0.031 (0.048)	-0.001 (0.076)	-0.029 (0.084)	0.002 (0.106)	-
Cancer $_{i,w}$	-0.041 (0.037)	-0.028 (0.062)	-0.063** (0.032)	-0.083 (0.061)	-0.098 (0.074)	-0.037 (0.096)	-
Lung $_{i,w}$	0.005 (0.032)	-0.025 (0.063)	-0.051 (0.033)	-0.119 (0.073)	-0.024 (0.042)	-0.034 (0.081)	-
Psych $_{i,w}$	-0.043 (0.031)	-0.073 (0.056)	0.018 (0.031)	0.006 (0.054)	-0.061 (0.049)	-0.042 (0.076)	-
Heart $_{i,w}$	-0.015 (0.031)	-0.005 (0.048)	-0.032 (0.028)	-0.037 (0.049)	0.015 (0.048)	0.073 (0.069)	-
Diab $_{i,w}$	0.044* (0.025)	0.051 (0.042)	0.023 (0.032)	0.041 (0.051)	-0.014*** (0.050)	-0.116 (0.076)	-
Arthri $_{i,w}$	-0.036** (0.017)	-0.061** (0.031)	0.001 (0.021)	-0.012 (0.037)	0.003 (0.035)	0.004 (0.063)	-
Married $_{i,w}$	-0.008 (0.019)	-0.035 (0.031)	-0.019 (0.021)	-0.022 (0.035)	-0.012 (0.034)	-0.076 (0.052)	+
Divorce $_{i,w}$	-0.006 (0.022)	-0.038 (0.037)	-0.032 (0.024)	-0.041 (0.041)	-0.017 (0.041)	-0.053 (0.062)	-
Widow $_{i,w}$	0.013 (0.025)	0.009 (0.038)	-0.031 (0.029)	-0.011 (0.044)	-0.027 (0.042)	-0.061 (0.064)	-
Hispanic $_{i}$	0.052*** (0.019)	0.082*** (0.031)	-0.023 (0.021)	-0.018 (0.036)	-0.055 (0.036)	0.037 (0.052)	0
Black $_{i}$	-0.012 (0.014)	0.007 (0.023)	0.018 (0.016)	0.035 (0.025)	0.017 (0.024)	0.031 (0.041)	0
Age $_{i,w}$	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.000 (0.002)	0.002 (0.002)	0.003 (0.002)	0
Male $_{i}$	-0.005 (0.009)	0.009 (0.015)	-0.012 (0.011)	-0.028 (0.018)	0.015 (0.018)	0.026 (0.033)	0
Grad $_{i}$	0.017* (0.009)	0.031* (0.017)	-0.022** (0.011)	-0.036* (0.021)	0.019 (0.017)	0.025 (0.033)	0
Protestant $_{i,w}$	0.000 (0.017)	0.027 (0.031)	-0.025 (0.018)	0.004 (0.034)	-0.026 (0.027)	-0.031 (0.048)	+
Catholic $_{i,w}$	-0.009 (0.017)	0.002 (0.003)	-0.026 (0.019)	-0.006 (0.036)	-0.005 (0.031)	-0.035 (0.054)	+
Jewish $_{i,w}$	0.001 (0.037)	0.058 (0.058)	0.013 (0.036)	0.059 (0.064)	-0.031 (0.061)	-0.101 (0.122)	+
Exercise $_{i,w}$	0.000 (0.008)	-0.008 (0.014)	-0.003 (0.009)	-0.003 (0.017)	0.012 (0.014)	-0.034 (0.027)	+
BMI $_{i,w}$	0.001 0.000	0.001 (0.001)	0.000 0.000	0.001 (0.001)	0.000 (0.001)	-0.001 (0.002)	-
Smoke $_{i,w}$	0.001 (0.012)	0.012 (0.019)	-0.013 (0.013)	0.001 (0.022)	0.055*** (0.019)	0.052 (0.033)	-
Earnings $_{i,w}$	-0.002** (0.001)	-0.006** (0.002)	0.006*** (0.002)	0.01*** (0.003)	-0.002 (0.002)	-0.005 (0.005)	+
Depressed $_{i,w}$	0.000 (0.015)	0.049* (0.026)	-0.021 (0.018)	-0.008 (0.028)	0.009 (0.031)	0.021 (0.051)	-
Happy $_{i,w}$	0.003 (0.015)	0.018 (0.026)	0.017 (0.018)	0.013 (0.029)	0.057** (0.028)	0.063 (0.048)	+
Lonely $_{i,w}$	-0.045*** (0.015)	-0.062** (0.025)	-0.002 (0.018)	0.012 (0.031)	0.037 (0.031)	0.062 (0.047)	-
DadAlive $_{i,w}$	0.009 (0.013)	0.013 (0.023)	-0.022 (0.015)	-0.027 (0.028)	0.004 (0.022)	-0.017 (0.041)	+
MomAlive $_{i,w}$	-0.004 (0.008)	-0.031** (0.015)	0.001 (0.011)	-0.003 (0.018)	-0.014 (0.016)	-0.044 (0.028)	+
Kids $_{i,w}$	0.001 (0.002)	0.003 (0.003)	-0.001 (0.002)	0.002 (0.004)	0.001 (0.003)	0.002 (0.005)	+
StressJob $_{i,w}$	0.007 (0.011)	0.011 (0.019)	-0.025** (0.013)	-0.052** (0.023)	0.012 (0.021)	0.001 (0.035)	-
PhysJob $_{i,w}$	0.003 (0.011)	0.017 (0.021)	0.009 (0.013)	0.004 (0.023)	0.015 (0.019)	0.048 (0.036)	-
Working $_{i,w}$	0.005 (0.012)	-0.002 (0.022)	-0.019 (0.014)	-0.018 (0.025)	-0.001 (0.022)	-0.009 (0.041)	+
Drink $_{i,w}$	0.009 (0.008)	-0.002 (0.014)	-0.015 (0.009)	-0.026 (0.016)	0.013 (0.015)	0.002 (0.026)	+
South $_{i,w}$	-0.005 (0.008)	-0.005 (0.014)	0.022 (0.009)	0.026 (0.016)	0.011 (0.015)	0.011 (0.026)	+
R-squared	0.254	0.012	0.253	0.015	0.265	0.021	
Observations	3995	3995	3039	3039	1400	1400	

Dependent variable is the change in the individual-reported expected probability of living to the age of 75 from wave to wave. For example "LE2 - LE1" represents the reported probability in wave 2 minus the reported probability in wave 1. Coefficients with robust standard errors in parenthesis reported. For Tobit regressions, Pseudo R-squared is reported. Constant not reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Although the number of significant predictors is minimal, the associations amongst these regressors are consistent with our theoretical expectations. Health shocks and other dynamic factors such as, the recent lonely state of individuals, are still observed to have negative effects on the changes of individuals' longevity expectations. Perhaps why this estimation does not provide strong evidence of the relationship between longevity expectations and mortality risks is simply the duration between HRS interview stages. The two year period in between waves leaves a significant lapse of time to become exposed to new risks or to experience positive and negative health shocks; thereby increasing the variability of individuals' longevity expectations. This observation is portrayed in our results in that the variances and robust standard errors of almost every single variable across all waves is larger in Model (3) compared to those of Model (2) for the same variables.

ii. Longevity Expectations' Relationship with Death

Results of the Probit estimation of Model (4) can be found in Table 5 on the following page.

Table 5. The Relationship between Longevity Expectations and Surviving to the Next Wave

Variable	Alive in Wave 2	Alive in Wave 3	Alive in Wave 4	Alive in Wave 5	Alive in Wave 6	Alive in Wave 7	Alive in Wave 8	Exp. Sign
$LE_{i,w-1}$	0.393*** (0.097)	0.562*** (0.105)	0.352*** (0.107)	0.527*** (0.117)	0.570*** (0.121)	0.625*** (0.159)	0.749*** (0.165)	+
$Smoke_{i,w-1}$	-0.252*** (0.064)	-0.327*** (0.068)	-0.394*** (0.075)	-0.388*** (0.078)	-0.333*** (0.081)	-0.217** (0.111)	-0.183 (0.113)	-
$Drink_{i,w-1}$	0.045 (0.063)	0.164** (0.067)	0.169** (0.074)	0.014 (0.076)	0.084 (0.076)	0.057 (0.101)	0.119 (0.104)	+
$Earnings_{i,w-1}$	0.063*** (0.019)	0.029* (0.016)	0.082*** (0.022)	0.091*** (0.024)	0.069*** (0.022)	0.014 (0.018)	0.051* (0.028)	+
$Hbp_{i,w-1}$	-0.161** (0.064)	0.188 (0.206)	0.192 (0.231)	0.218 (0.214)	-0.228 (0.145)	-0.004 (0.200)	-0.045 (0.187)	-
$Stroke_{i,w-1}$	-0.323*** (0.116)	0.455 (0.497)	0.321 (0.464)	-0.629** (0.271)	-0.905*** (0.195)	-0.967*** (0.226)	-0.451 (0.295)	-
$Cancer_{i,w-1}$	-0.705*** (0.084)	-0.984*** (0.154)	-0.913*** (0.173)	-0.772*** (0.175)	-0.899*** (0.162)	-0.978*** (0.188)	-0.989*** (0.199)	-
$Lung_{i,w-1}$	-0.149* (0.089)	-0.215 (0.205)	-0.768*** (0.216)	-0.439** (0.200)	-0.399* (0.237)	0.285 (0.435)	-0.422 (0.258)	-
$Psych_{i,w-1}$	0.152 (0.095)	-0.529*** (0.148)	-0.133 (0.236)	-0.304 (0.195)	0.067 (0.255)	-0.093 (0.282)	-0.162 (0.275)	-
$Heart_{i,w-1}$	-0.387*** (0.073)	-0.231 (0.171)	-0.512*** (0.169)	-0.362** (0.182)	-0.106 (0.208)	-0.561*** (0.195)	0.081 (0.288)	-
$Diab_{i,w-1}$	-0.292*** (0.077)	-0.402** (0.177)	-0.031 (0.243)	-0.449** (0.201)	-0.010 (0.228)	-0.218 (0.225)	-0.181 (0.238)	-
$Arthri_{i,w-1}$	0.046 (0.064)	-0.256** (0.115)	-0.226* (0.133)	-0.031 (0.165)	0.254 (0.213)	-0.284* (0.163)	-0.151 (0.193)	-
Constant	2.13 (0.095)	1.863 (0.078)	1.935 (0.084)	1.833 (0.089)	1.68 (0.093)	1.93 (0.119)	1.602 (0.119)	
Pseudo R-squared	0.119	0.091	0.090	0.095	0.094	0.104	0.101	
Observations	11633	9638	8396	6958	5937	4628	3466	

Dependent variable is a 1 if the individual was alive in the wave and 0 otherwise. Probit coefficients with robust standard errors in parenthesis reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

The results from Table 5 illustrate that total earnings from the last wave are associated with being more likely to survive to the next wave, while major health shocks and smoking status in the preceding wave are indicative of a lower probability of surviving to the next wave, with the diagnosis of cancer being the most pernicious of the regressors. Moreover, our variable of interest, $LE_{i,w-1}$, is positive and significant in every wave with an average positive association of 0.540. This provides strong support for this paper's second hypothesis that individuals' longevity expectations can signal surviving to the next wave, and in these estimations, longevity expectations' association, on average, explains over 50% of such a relationship.

iii. Testing Aggregate Rationality

Testing Model (7) for rationality in the sample is segmented into three different parts with each focusing on a specific aspect of The Theory of Rational Expectations. First, systematic biases are examined by only regressing forecast errors on the constant bias term, $\varphi_{i,w}$, The results for the OLS, Tobit, and Heckman Two-Step estimations can be found in Table 6 below and the first step of the Heckman Selection Model is shown in Appendix 1B.

Table 6. Rationality Test using Only a Constant Term

Variable	OLS	Tobit	Heckman Two-Step
$\varphi_{i,w}$ (Constant)	0.079*** (0.012)	0.075*** (0.012)	0.051* (0.027)
Observations	1670	1670	1796
Censored Observations			137

Dependent variable is the whether or not the individual was alive at 75 (1 if yes, 0 if no) minus the individual-reported expected probability of living to the age of 75 in wave 1. Coefficients with robust standard errors in parenthesis reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

The positive statistical significance of $\varphi_{i,w}$ is associated with the premise that individuals on average underestimate their chances of living to age 75. Therefore, these results suggest that under the assumption that individuals anticipate the probability of experiencing health shocks, than individuals within the sample exhibit irrationality in the form of a ‘frail’ or ‘pessimistic’ outlook on life wherein individuals do not expect to live as long as they actually do.

With the addition of $\lambda_{i,T\text{-}w}$ to introduce different forms of news into Model (7), testing for rationality becomes unclear. As stated in the Rationality Hypothesis earlier, health shocks and other forms of news should not be incorporated into individuals’ longevity expectations. However, in the case of health information, individuals may be able to anticipate contracting a certain disease or ailment to a certain degree. Further, individuals may be able to estimate a probability on contracting the health condition based upon that anticipation and integrate it into their longevity expectation. Therefore, to interpret $\varphi_{i,w}$ in this case as a bias term would likely be incorrect[30]; however, we still regress Model (7) with $\varphi_{i,w}$ and $\lambda_{i,T\text{-}w}$ as regressors to examine how individuals' longevity expectations are associated with health shocks. The results from the OLS, Tobit, and Heckman Two-Step estimations are in Table 7 on the next page with the first step of the Heckman Selection Model again shown in Appendix 1B.

Table 7. Rationality Test using Constant Term and Health Shocks

Variable	OLS	Tobit	Heckman Two-Step
$\varphi_{i,w}$ (Constant)	0.029 (0.017)	0.024 (0.019)	0.012 (0.029)
$\lambda_{i,T\text{-}w}$			
$Hbp_{i,T\text{-}w}$	0.085*** (0.028)	0.089*** (0.031)	0.085*** (0.029)
$Stroke_{i,T\text{-}w}$	-0.016 (0.042)	-0.017 (0.047)	-0.003 (0.042)
$Cancer_{i,T\text{-}w}$	0.006 (0.035)	0.008 (0.037)	0.011 (0.033)
$Lung_{i,T\text{-}w}$	-0.078 (0.047)	-0.089* (0.052)	-0.073* (0.042)
$Psych_{i,T\text{-}w}$	-0.021 (0.054)	-0.022 (0.058)	-0.010 (0.048)
$Heart_{i,T\text{-}w}$	0.072** (0.029)	0.072** (0.031)	0.066** (0.030)
$Diab_{i,T\text{-}w}$	0.029 (0.037)	0.029 (0.039)	0.041 (0.036)
$Arthri_{i,T\text{-}w}$	0.098*** (0.027)	0.105*** (0.029)	0.094*** (0.028)
Observations	1670	1670	1796
Censored Observations			137

Dependent variable is the whether or not the individual was alive at 75 (1 if yes, 0 if no) minus the individual-reported expected probability of living to the age of 75 in wave 1. Coefficients with robust standard errors in parenthesis reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Again, note that while $\varphi_{i,w}$ is significant in this regression, it does not serve as an accurate measure of aggregate rationality, because individuals may be able to partially anticipate health shocks and incorporate the probability of experiencing them into their longevity expectations.

[30] Appendices 2A through 2C outlines an numerical example of this inconsistency

Therefore, the degree to which one assumes individuals can anticipate a health shock determines the resulting rationality conclusion of the last two estimates of Model (7). If one is to assume full anticipation then we find a slight pessimism bias, if one is to assume health shocks are not anticipatory then we cannot reject rationality within the sample. Nonetheless, one cannot ignore the statistical significance of health shocks on individuals' forecast errors.

The relevant inferences to be drawn from Table 7's estimation of Model (7) in relation to health shocks have direct connections to the forecast errors themselves. The negatively associated health shock, lung disease, can be interpreted as follows: if an individual has had lung disease between wave 2 and turning age 75, then he is more likely to have incorrectly overestimated his survival chances at age 75. This result can be examined in two ways: either the event was a complete shock and the downtown of his health was the cause for the overestimation in Wave 1, or the individual did not anticipate contracting lung disease, or at least not as likely as he should have calculated. The positively associated health shocks, high-blood pressure, heart disease, and arthritis, can be interpreted in the same manner; only inversely.

Lastly, Model (7) is now tested in its full form, but once again, the constant bias term, $\varphi_{i,w}$, is not a clear measure for judging rationality unless it is jointly insignificant with β_3[31]. The results of the full version of Model (7) are shown in Table 8 on the next page with the first step of the Heckman Selection Model following in Appendix 1B.

[31] Appendices 2A through 2C outlines a numerical example of this inconsistency

Table 8. Rationality Test using Constant Term, Health Shocks, and Known Information in Wave 1

Variable	OLS	Tobit	Heckman Two-Step
$\varphi_{i,w}$ (Constant)	0.001 (0.269)	-0.007 (0.284)	-0.407 (0.424)
$Hbp_{i,T-w}$	0.079*** (0.029)	0.082*** (0.031)	0.072** (0.029)
$Stroke_{i,T-w}$	0.024 (0.042)	0.025 (0.044)	0.025 (0.041)
$Cancer_{i,T-w}$	0.022 (0.035)	0.026 (0.038)	0.023 (0.033)
$Lung_{i,T-w}$	-0.012 (0.049)	-0.018 (0.053)	-0.013 (0.043)
$Psych_{i,T-w}$	-0.015 (0.053)	-0.018 (0.056)	-0.009 (0.048)
$Heart_{i,T-w}$	0.064** (0.029)	0.065** (0.031)	0.061** (0.031)
$Diab_{i,T-w}$	0.043 (0.036)	0.046 (0.038)	0.046 (0.036)
$Arthri_{i,T-w}$	0.086*** (0.028)	0.094*** (0.029)	0.091*** (0.028)
$Married_{i,w}$	-0.032 (0.069)	-0.034 (0.076)	-0.038 (0.071)
$Divorce_{i,w}$	-0.133 (0.086)	-0.137 (0.094)	-0.125 (0.086)
$Widow_{i,w}$	-0.172* (0.092)	-0.174* (0.099)	-0.215** (0.095)
$Hispanic_{i}$	0.041 (0.049)	0.048 (0.053)	0.033 (0.052)
$Black_{i}$	-0.069* (0.038)	-0.083** (0.041)	-0.062 (0.038)
$Age_{i,w}$	0.004 (0.004)	0.004 (0.004)	0.007 (0.005)
$Male_{i}$	-0.082*** (0.029)	-0.084*** (0.032)	-0.090*** (0.031)
$Grad_{i}$	-0.045 (0.030)	-0.048 (0.031)	-0.041 (0.035)
$Protestant_{i,w}$	-0.062 (0.059)	-0.071 (0.063)	-0.041 (0.065)
$Catholic_{i,w}$	-0.019 (0.061)	-0.026 (0.066)	-0.003 (0.068)
$Jewish_{i,w}$	-0.163 (0.103)	-0.173 (0.107)	-0.133 (0.115)
$Exercise_{i,w}$	0.032 (0.031)	0.034 (0.033)	0.042 (0.034)
$BMI_{i,w}$	0.001 (0.003)	0.001 (0.003)	0.002 (0.003)
$Smoke_{i,w}$	-0.120*** (0.033)	-0.131*** (0.035)	-0.109*** (0.032)
$Earnings_{i,w}$	0.008** (0.003)	0.007** (0.003)	0.001** 0.000
$Depressed_{i,w}$	0.000 (0.032)	0.006 (0.034)	0.012 (0.033)
$Happy_{i,w}$	-0.125 (0.087)	-0.121 (0.089)	-0.103 (0.089)
$Lonely_{i,w}$	0.057* (0.034)	0.069* (0.035)	0.059* (0.036)
$DadAlive_{i,w}$	0.011 (0.054)	0.005 (0.057)	0.009 (0.057)
$MomAlive_{i,w}$	-0.032 (0.028)	-0.036 (0.031)	-0.029 (0.029)
$Kids_{i,w}$	-0.002 (0.005)	-0.001 (0.005)	0.003 (0.006)
$StressJob_{i,w}$	0.000 (0.033)	0.003 (0.035)	0.008 (0.037)
$PhysJob_{i,w}$	0.045 (0.036)	0.049 (0.038)	0.069 (0.044)
$Working_{i,w}$	0.025 (0.039)	0.022 (0.041)	0.004 (0.046)
$Drink_{i,w}$	0.065** (0.026)	0.064** (0.028)	0.074*** (0.027)
$South_{i,w}$	-0.026 (0.026)	-0.029 (0.027)	-0.027 (0.026)
Observations	1659	1659	1796
Censored Observations			137

Dependent variable is the whether or not the individual was alive at 75 (1 if yes, 0 if no) minus the individual-reported expected probability of living to the age of 75 in wave 1. Coefficients with robust standard errors in parenthesis reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Because we do not observe joint insignificance between $\varphi_{i,w}$ and the vector affiliated with β_3, a clear test of rationality is unable to be supported. However, the results do suggest that we can attribute forecast errors to particular variables. First, we recognize that health shocks behave in the same manner in this more comprehensive specification. We see a predominant underestimation in individuals' longevity expectations wherein the individual become subject to health shocks before they reach age 75. A simple interpretation of these positive associations suggests that individuals overestimate their chances of contracting in the first place or surviving with high-blood pressure, heart disease, and arthritis to the age of 75.

More intriguingly, yet perhaps all too predictably, the results from the estimation of Model (7)'s full version suggest that men and smokers are, on average, more likely to exhibit 'invincibility bias,' meaning they expect to live longer than they actually do. This overestimation could be attributable to an innate bias amongst men and smokers to expect to overcome the risks that they, on average, face, or the overestimation could perhaps stem from a miscalculation of just how much mortality risk is attached to the average male and average smoker.

Lastly, a Probit form of Model (7) was employed to generate a forecast estimate for each individual based upon the information known within the model.[32] These predicted longevity expectations were then compared to individuals' actual longevity expectations in determining the accuracy of individuals' survival predictions for the horizon age of 75. The distribution of each individual's forecasts error is show in Figure 1 on the next page.

---

[32] $$\hat{Alive75}_{i,T-w} = \alpha + \beta_1 Married_{i,w} + \beta_2 Divorced_{i,w} + \beta_3 Widow_{i,w} + \beta_4 Hispanic_i + \beta_5 Black_i + \beta_6 Age_{i,w} + \beta_7 Male_{i,} + \beta_8 Grad_i + \beta_9 Protestant_{i,w} + \beta_{10} Catholic_{i,w} + \beta_{11} Jewish_{i,w} + \beta_{12} Exercise_{i,w} + \beta_{13} BMI_{i,w} + \beta_{14} Smoke_{i,w} + \beta_{15} Earnings_{i,w} + \beta_{16} Depressed_{i,w} + \beta_{17} Happy_{i,w} + \beta_{18} Lonely_{i,w} + \beta_{19} DadAlive_{i,w} + \beta_{20} MomAlive_{i,w} + \beta_{21} \#Kids_{i,w} + \beta_{22} StressJob_{i,w} + \beta_{23} PhysJob_{i,w} + \beta_{24} Working_{i,w} + \beta_{25} Drink_{i,w} + \beta_{26} South_{i,w} + \beta_{27} Hbp_{i,T-w} \beta_{28} Stroke_{i,T-w} + \beta_{29} Cancer_{i,T-w} + \beta_{30} Lung_{i,T-w} + \beta_{31} Psych_{i,T-w} + \beta_{32} Heart_{i,T-w} + \beta_{33} Diab_{i,T-w} + \beta_{34} Arthri_{i,T-w} + \varepsilon_i$$

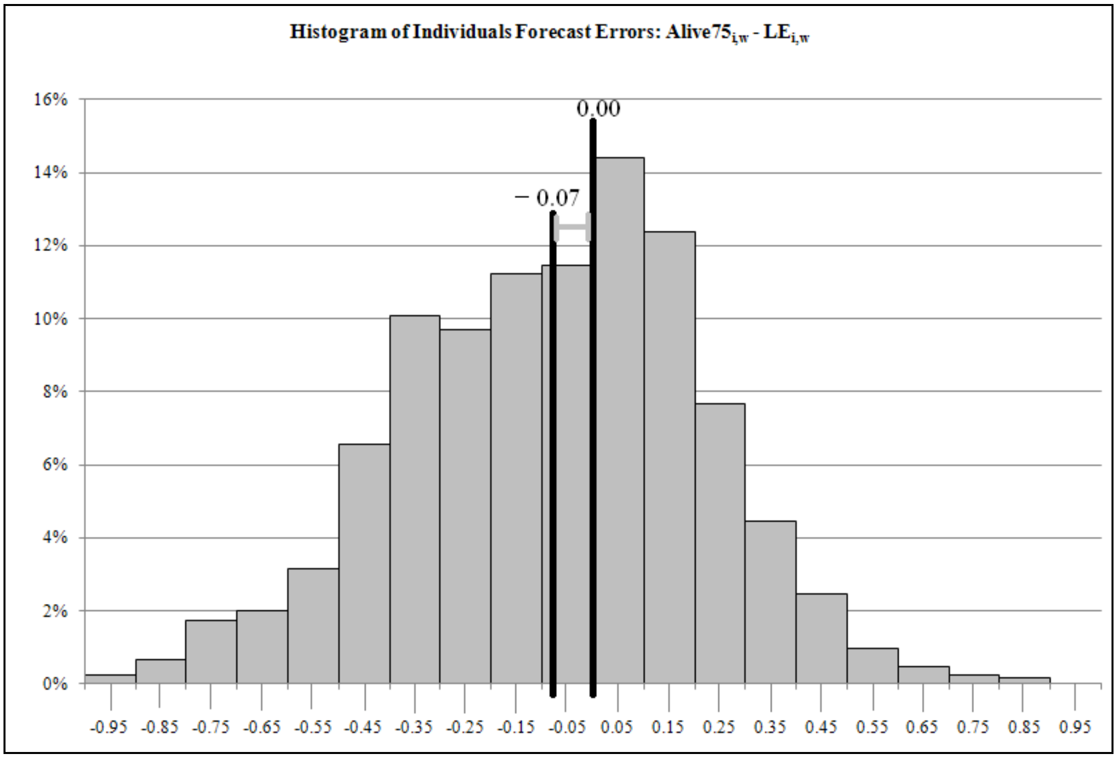

Figure 1. Histogram of Individual Forecast Errors

Figure 1 illustrates the slight pessimism bias aforementioned. If individuals' longevity expectations were on average rational, then we would see the distribution's mean exist at 0, instead of the calculated mean diagramed above at -0.07. This is also what we roughly observed in the earlier OLS, Tobit, and Heckman regressions.

This accuracy test of forecasting errors is furthered through calculating the mean deviation between the Probit generated estimates' and individuals' own longevity expectations. The resulting mean deviation along with a further breakdown of the mean deviations of individuals who ended up surviving to age 75 versus those who did not are shown in Table 9 on the next page.

Table 9. Mean Deviation of Individual Forecast Errors

Survival Status	Mean Deviation
Entire sample of individuals with known survivor status at age 75	24.497 percentage points
Alive at age 75	23.436 percentage points***
Dead at age 75	27.252 percentage points***

***A hypothesis test for two means reveals that the mean deviation for the group that is ultimately alive at 75 is statistically different from the mean deviation for the group that is ultimately dead at 75 at a 0.001 level of significance.

The contents of the Table 9 illustrate that individuals' longevity expectations on average differ from what information pertaining to them would predict by nearly 25 percentage points. Although, there is no standard or instrument for judging this accuracy test, we can examine the inter-sample accuracy of our two main groups of interest, Alive at 75 and Dead at 75. Again, through calculating the mean deviations for each of the observations in each group, the results suggest that individuals who survived to age 75 more closely estimated their longevity forecasts than those individuals who did not survive.[33]

V. Implications

This paper answered three research questions pertaining to individuals' longevity expectations: Are they indicative of the actual mortality risk individuals personally face? Do they have a predictive power in signaling surviving to age 75? And lastly, to what degree are they

[33] Independent Groups t-Test for differences in means

Group	Sample Size	Mean	Standard Deviation
Dead at 75	433	27.252	0.196
Alive at 75	1226	23.436	0.182
t-Test Results			
Test Statistic (distributed t)	3.562	Degrees of Freedom	721.55

unbiased and accurate? The answers to these questions contain insight that is directly meaningful for personal, social, economical, and bureaucratic ends.

The chief conclusions are that although individuals' longevity expectations are reflective of mortality risk and predictive of one's survival status; ultimately however, if one assumes that individuals fully anticipate health shocks or the probability of experiencing a health shock, then individuals do exhibit a slight pessimism bias in estimating how long they expect to live. These conclusions suggest the possibility of irrationality within the sample; calling into question the ability of individuals to realistically optimize their life-cycle decision plans, which holds direct implications for all multi-period models and for insurance, financial, and government planning. The answer to "how accurate" depends on what one considers accurate. The average "error" between individual longevity expectations and econometrically calculated longevity expectations is about 25 percentage points.

VI. Extensions for Future Research

Future research regarding individuals' longevity expectations should begin by expanding the sample analyzed. The HRS dataset proved to be extremely useful; however, the limited age range provided a narrow scope. Researchers should explore the possibility of gathering longevity expectations of individuals from a wider range of ages even if the longitudinal tracking of those individuals to the horizon age is lengthy. At this point, the AHEAD (Asset and Health Dynamics Amongst the Oldest Old) dataset is also available, which studies individuals at a target age of 81 and older. This dataset contains similar questions, but has a slightly different question inquiring into individuals' longevity expectations; it asks individuals about their chances of living to age 85 or ten more years.

Aside from expanding the dataset, researchers may also want to target the longevity expectations of specific groups. This paper has already established that some of the possible determinants of bias within the sample could be the experience of certain health shocks, being male, and smoking. Picone and Sloan (2003) and Schoenbaum (1997) explore the topic of smoking, but do not explicitly test for rationality or any degree accuracy.

Lastly, adding additional variables will always be practical, as one of the premises of establishing a rational expectation is to consider all relevant information. A researcher would always be well-served to include as many related variables as is practical in examining one's longevity expectation, in order to control for the impact of different types of known information and news.

VII. Conclusion

The purpose of this analysis was to examine individuals' subjective probabilities of surviving to the age of 75. The preliminary theological framework of Hamermesh (1985), Hurd and McGarry (2002) and Smith et al. (2001) provided the foundation of the paper; however the econometric work of Davies and Lahiri (1995 and 1999) was ultimately employed to test the Theory of Rational Expectations within the sample.

The model's preliminary specifications show that these subjective probabilities of living to the age of 75, termed "longevity expectations," are consistent with the actual mortality risks to which individuals are personally exposed. Further, the model also finds that individuals' longevity expectations are significant predictors of an individual's survival in the following period, thereby demonstrating that an individual's longevity expectation is predictive of one's death; a proposition that could be potentially useful in any context concerning life-cycle decisions and behavior.

Later specifications of the model find evidence that if the assumption is made that individuals can fully or partially anticipate health shocks and incorporate the shocks' effects into their longevity expectations, then individuals exhibit marginally underestimated, yet statistically biased forecast errors in predicting their survival statuses at age 75. This finding is inconsistent with the Theory of Rational Expectations outlined in earlier portions of the paper. The results indicate that this slight pessimism in the formulation of individuals' longevity expectations of surviving to age 75 is between 5 to 7 percentage points depending on the regression type. This conclusion however, is dependent upon the assumption that future health shocks are incorporated in longevity expectations (or more accurately that the probability of undertaking a future health shock is incorporated.)

In addition to testing rationality, the paper also seeks to measure the relative accuracy of the forecasts. The results show that individuals miscalculate their probabilities of surviving to the age of 75 by an average of nearly 25 percentage points with survivors' at 75 estimates being comparatively better estimated than the deceased's. Individuals who end up surviving at age 75 more correctly estimate their survival chances than those who expire before age 75. Therefore, in addressing the question, how well can individuals predict their own death? The results reply, individuals can adequately enough, but better if you are a future survivor.

This evidence of possible irrationality and severely miscalculated forecast estimates entails marked implications for individuals' personal planning, such as consumption, investing, and retirement decisions, the finance and insurance industries' profit-seeking and hedging strategies, and the government's planning of social programs whereby these entities will have to account for biases and grave miscalculations in individuals' forecast estimates.

VIII. References

Ando, A. and F. Modigliani. 1963. "The 'Life Cycle' Hypothesis of Saving: Aggregate Implications and Tests." *American Economic Review*, 53(1): 55-84.

Arthur, W. 1981. "The Economics of Risks to Life." *American Economic Review,* 71(1): 54-64.

Barro, R. 1977. "Unanticipated Money Growth and Unemployment in the United States." *American Economic Review*, 67(2): 101-115.

Barro, R. 1978. "Unanticipated Money, Output, and the Price Level in the United States." *American Economic Review*, 86(3): 549-580.

Benitez-Silva, H. and D. Dwyer. 2005. "The Rationality of Retirement Expectations and the Role of New Information." *Review of Economics and Statistics*, 87(3): 587-592.

Benitez-Silva, H., D. Dwyer, W. Gayle, and T. Muench. 2008. "Expectations in Micro Data: Rationality Revisited." *Empirical Economics*, 34(2): 341-416.

Davies, A. and K. Lahiri. 1995. "A New Framework for Testing Rationality and Measuring Aggregate Shocks Using Panel Data." *Journal of Econometrics*, 68(1): 205-227.

Davies, A. and K. Lahiri. 1999. "Re-examining the Rational Expectations Hypothesis Using Panel Data on Multi-Period Forecasts." In Analysis of Panels and Limited Dependent Variable Models, ed. Hsiao, C., M. Peseran, K. Lahiri, and F. Lung. Cambridge: Cambridge University Press.

Fisher, I. 1930. The Theory of Interest. New York: Augustus M. Kelley Publishers.

Friedman, M. 1957. A Theory of the Consumption Function. Princeton: Princeton University Press.

Frydman, R. and P. Rappoport. 1987. "Is the Distinction between Anticipated and Unanticipated Money Growth Relevant in Explaining Aggregate Output?" *American Economic Review,* 77(4): 693-703.

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

Harrod, R. 1948. Towards a Dynamic Economics: Some Recent Developments of Economic Theory and Their Application to Policy. London: Macmillan Publishers.

Heady, E. and D. Kaldor. 1954. "Expectations and Errors in Forecasting Agricultural Prices." *Journal of Political Economy,* 62(1): 34-47.

Heckman, J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica*, 47(1): 153-161.

HRS Dataset 1992-2006. RAND Corporation. Center for the Study of Aging Founded by the National Institute on Aging. File Type: RAND intercooled STATA file < http://hrsonline.isr.umich.edu/ >.

Hurd, M. and K. McGarry. 2002. “The Predictive Validity of Subjective Probabilities of Survival.” *The Economic Journal*, 112(482): 966-985.

Hurd, M., D. McFadden, and A. Merrill. 2001. “Predictors of Mortality Among the Elderly.” In Themes in the Economics of Aging, ed. Wise, D. Chicago: University of Chicago Press.

Juster, F. and R. Suzman. 1995. “An Overview of the Health and Retirement Study.” *The Journal of Human Resources*, 30(5): S7-S56.

Lucas, R. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4(2): 103–124.

Lucas, R. 1977. “Some International Evidence on Output-Inflation Tradeoffs: Reply.” *American Economic Review*, 67(4): 731-756.

Machina, M. 1987. “Choice under Uncertainty: Problems Solved and Unsolved.” *Journal of Economic Perspectives*, 1(1): 121-154.

Mishkin, F. 1982a. “Does Anticipated Aggregate Demand Policy Matter? An Econometric Investigation.” *Journal of Political Economy*, 90(1): 22-51.

Mishkin, F. 1982b. “Does Anticipated Aggregate Demand Policy Matter? Further Econometric Results.” *American Economic Review*, 72(4): 788-802.

Modligliani, F. and H. Weingartner. 1958. “Forecasting Uses of Anticipatory Data on Investment and Sales.” *Quarterly Journal of Economics*, 72(1): 23-54.

Muth, J. 1961. “Rational Expectations and the Theory of Price Movements.” *Econometrica*, 29(1): 315-335.

Picone, G. and F. Sloan. 2003. “Smoking Cessation and Lifestyle Changes.” In Frontiers in Health Policy Research, Volume 6, ed. Cutler, D. and A. Garber. Cambridge: MIT Press.

Schoenbaum, M. 1997. “Do Smokers Understand the Mortality Effects of Smoking? Evidence from the Health and Retirement Survey.” *American Journal of Public Health*, 87(5): 755-759.

Smith, V., D. Taylor, and F. Sloan. 2001. “Longevity Expectations and Death: Can People Predict Their Own Demise?” *American Economic Review*, 91(4): 1126-1134.

Thaler, R. and S. Rosen. 1976. "The Value of Saving a Life: Evidence from the Labor Market." In Household Production and Consumption, ed. Terleckyj, N. New York: Columbia University Press.

Viscusi, W., J. Hakes, and A. Carlin. 1997. "Measures of Mortality Risks." *Journal of Risk and Uncertainty*, 14(3): 213-233.

Viscusi, W. 1979. Employment Hazards: An Investigation of Market Performance. Cambridge: Harvard University Press.

White, H. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica*, 50(1): 1-25.

Wilson, R. 1979. "Analyzing the Daily Risks of Life." *Technology Review*, 81(4): 41-47.

Appendix 1A: Correction for Selection Bias

Attrition in the 1,810 individual sample (used to test the rationality hypothesis) was explored by running a Probit regression on individuals who were not marked as dead or alive at the time they were scheduled to turn age 75. This dependent variable was then comprised of 1's for individuals who were counted as EITHER dead or alive at 75 and 0 otherwise –i.e. dropped from sample, unaccounted for, or their whereabouts were uncertain. This Probit regression yielded 5 regressors that were significant in distinguishing individuals that made it into the sample and those that were subject to attrition. Those independent variables were: Age in Wave 1, prior history with cancer, prior history with stroke, working at a physically intensive job in Wave 1, and the number of children the respondent has living. Therefore, the first step of the Heckman Two-Step procedure was enacted. Essentially, a Probit model is employed with the aforementioned '*dead or alive at 75 =1 variable and 0 if subject to attrition'* variable as the regression's dependent term. (The independent variables are shown in the following appendix, Appendix 1B) The corrective term produced from this regression is lambda (also referred to as a Mills ratio)—also shown in Appendix 1B. Lambda is calculated by taking the probability density function of the positive outcome of the Probit model and dividing it by the model's cumulative density function. This lambda term is then included in the main regression to account for the probability of an observation making it into the sample in the first stage.

Appendix 1B: Selection Models (for Heckman Two-Step Process in Tables 6, 7, and 8)

Selection Stage for Heckman Two-Step Model in Table 6

Dependent Variable = 1 if known status at 75 (alive or dead); 0 if censored (i.e. dropped from sample).			
$Hbp_{i,w}$	0.105 (0.097)	$Jewish_{i,w}$	0.308 (0.415)
$Stroke_{i,w}$	0.556* (0.315)	$Exercise_{i,w}$	0.156 (0.128)
$Cancer_{i,w}$	-0.304* (0.170)	$BMI_{i,w}$	0.005 (0.011)
$Lung_{i,w}$	0.159 (0.171)	$Smoke_{i,w}$	0.117 (0.116)
$Psych_{i,w}$	-0.137 (0.173)	$Earnings_{i,w}$	0.033 (0.024)
$Heart_{i,w}$	0.231* (0.129)	$Depressed_{i,w}$	0.122 (0.128)
$Diab_{i,w}$	-0.024 (0.140)	$Happy_{i,w}$	0.182 (0.298)
$Arthri_{i,w}$	-0.064 (0.092)	$Lonely_{i,w}$	0.021 (0.108)
$Married_{i,w}$	-0.135 (0.274)	$DadAlive_{i,w}$	0.004 (0.208)
$Divorce_{i,w}$	0.012 (0.335)	$MomAlive_{i,w}$	0.021 (0.108)
$Widow_{i,w}$	-0.449 (0.332)	$Kids_{i,w}$	0.058** (0.022)
$Hispanic_{i}$	-0.045 (0.186)	$StressJob_{i,w}$	0.084 (0.134)
$Black_{i}$	0.059 (0.150)	$PhysJob_{i,w}$	0.278* (0.143)
$Age_{i,w}$	0.055** (0.019)	$Working_{i,w}$	-0.213 (0.152)
$Male_{i}$	-0.145 (0.122)	$Drink_{i,w}$	0.121 (0.097)
$Grad_{i}$	0.028 (0.129)	$South_{i,w}$	-0.021 (0.096)
$Protestant_{i,w}$	0.228 (0.215)	Constant	-0.277 (1.310)
$Catholic_{i,w}$	0.162 (0.226)	*lambda*	0.188 (0.257)
Observations	1796	*rho*	0.382
Censored Observations	137		

Coefficients with robust standard errors in parenthesis reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Selection Stage for Heckman Two-Step Model in Table 7

Dependent Variable = 1 if known status at 75 (alive or dead); 0 if censored (i.e. dropped from sample).			
$Hbp_{i,w}$	0.105 (0.097)	$Jewish_{i,w}$	0.308 (0.412)
$Stroke_{i,w}$	0.556* (0.314)	$Exercise_{i,w}$	0.156 (0.128)
$Cancer_{i,w}$	-0.304* (0.171)	$BMI_{i,w}$	0.005 (0.011)
$Lung_{i,w}$	0.159 (0.171)	$Smoke_{i,w}$	0.117 (0.116)
$Psych_{i,w}$	-0.137 (0.274)	$Earnings_{i,w}$	0.033 (0.024)
$Heart_{i,w}$	0.231* (0.129)	$Depressed_{i,w}$	0.122 (0.128)
$Diab_{i,w}$	-0.024 (0.141)	$Happy_{i,w}$	0.182 (0.295)
$Arthri_{i,w}$	-0.065 (0.092)	$Lonely_{i,w}$	0.022 (0.137)
$Married_{i,w}$	-0.135 (0.274)	$DadAlive_{i,w}$	0.004 (0.208)
$Divorce_{i,w}$	0.011 (0.335)	$MomAlive_{i,w}$	0.023 (0.111)
$Widow_{i,w}$	-0.449 (0.322)	$Kids_{i,w}$	0.058*** (0.022)
$Hispanic_{i}$	-0.045 (0.186)	$StressJob_{i,w}$	0.084 (0.134)
$Black_{i}$	0.059 (0.150)	$PhysJob_{i,w}$	0.278* (0.143)
$Age_{i,w}$	0.055*** (0.019)	$Working_{i,w}$	-0.213 (0.152)
$Male_{i}$	-0.145 (0.112)	$Drink_{i,w}$	0.121 (0.097)
$Grad_{i}$	0.028 (0.129)	$South_{i,w}$	-0.021 (0.096)
$Protestant_{i,w}$	0.228 (0.215)	Constant	-2.77 (1.313)
$Catholic_{i,w}$	0.162 (0.226)	*lambda*	0.103 (0.165)
Observations	1796	*rho*	0.214
Censored Observations	137		

Coefficients with robust standard errors in parenthesis reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Selection Stage for Heckman Two-Step Model in Table 8

Dependent Variable = 1 if known status at 75 (alive or dead); 0 if censored (i.e. dropped from sample).			
Hbp $_{i,w}$	0.105 (0.097)	*Jewish* $_{i,w}$	0.308 (0.415)
Stroke $_{i,w}$	0.556* 0.314	*Exercise* $_{i,w}$	0.156 0.128
Cancer $_{i,w}$	-0.304* (0.171)	*BMI* $_{i,w}$	0.005 (0.011)
Lung $_{i,w}$	0.159 0.171	*Smoke* $_{i,w}$	0.117 0.116
Psych $_{i,w}$	-0.137 (0.173)	*Earnings* $_{i,w}$	0.033 (0.024)
Heart $_{i,w}$	0.231* 0.129	*Depressed* $_{i,w}$	0.122 0.128
Diab $_{i,w}$	-0.024 (0.141)	*Happy* $_{i,w}$	0.181 (0.299)
Arthri $_{i,w}$	-0.064 0.092	*Lonely* $_{i,w}$	0.022 0.136
Married $_{i,w}$	-0.135 (0.274)	*DadAlive* $_{i,w}$	0.004 (0.208)
Divorce $_{i,w}$	0.011 0.335	*MomAlive* $_{i,w}$	0.021 0.108
Widow $_{i,w}$	-0.449 (0.322)	*Kids* $_{i,w}$	0.058*** (0.022)
Hispanic $_{i}$	-0.045 0.186	*StressJob* $_{i,w}$	0.084 0.134
Black $_{i}$	0.059 (0.150)	*PhysJob* $_{i,w}$	0.278* (0.143)
Age $_{i,w}$	0.055*** 0.019	*Working* $_{i,w}$	-0.213 0.152
Male $_{i}$	-0.145 (0.112)	*Drink* $_{i,w}$	0.121 (0.097)
Grad $_{i}$	0.028 0.129	*South* $_{i,w}$	-0.021 0.096
Protestant $_{i,w}$	0.228 (0.215)	Constant	-2.77 (1.313)
Catholic $_{i,w}$	0.162 0.226	*lambda*	0.413 (0.283)
Observations	1796	*rho*	0.812
Censored Observations	137		

Coefficients with robust standard errors in parenthesis reported. ***Significant at 0.01, **Significant at 0.05, *Significant at 0.1.

Appendix 2A: Numerical Example of Rationality Tests – Base Case (Ignoring Shocks and Known Information)

The purpose of the following examples is to illustrate simple numerical examples that show rational longevity expectations, optimistic longevity expectations, and pessimistic longevity expectations, and how those correspond to regression results.

1. Rational Longevity Expectations: Assume that everyone (correctly) anticipates that there is a 50% chance of being alive at 75 and a 50% chance of not being alive at 75. If these probabilities are correct and there are 4 individuals, then the number that are actually alive at 75 should be 2. So the data should look like the following.

Alive	LE	ForecastError
0	0.5	-0.5
0	0.5	-0.5
1	0.5	0.5
1	0.5	0.5

Notice that when everyone correctly anticipates the chance of being alive at 75, the average forecast error is 0, such that regressing Alive-LE on a constant results in the constant being 0.

2. Irrational (Optimistic) Longevity Expectations: Assume that everyone (incorrectly) anticipates that there is a 50% chance of being alive at 75, when there is actually only a 25% chance (such that individuals are overly optimistic.) If these probabilities are correct and there are 4 individuals, then the number that are actually alive at 75 should be 1. So the data should look like the following.

Alive	LE	ForecastError
0	0.5	-0.5
0	0.5	-0.5
0	0.5	-0.5
1	0.5	0.5

Note that when everyone is overly optimistic, the average forecast error is -0.25, such that regressing Alive-LE on a constant results in the constant being -0.25.

3. Irrational (Pessimistic) Longevity Expectations: Assume that everyone (incorrectly) anticipates that there is a 50% chance of being alive at 75, when there is actually a 75% chance (such that individuals are overly pessimistic.) If these probabilities are correct and there are 4 individuals, then the number that are actually alive at 75 should be 3. So the data should look like the following.

Alive	LE	ForecastError
0	0.5	-0.5
1	0.5	0.5
1	0.5	0.5
1	0.5	0.5

Note that when everyone is overly pessimistic, the average forecast error is 0.25, such that regressing Alive-LE on a constant results in the constant being 0.25.

Appendix 2B: Numerical Example of Rationality Tests – Incorporating Shocks

The purpose of the following examples is to illustrate that it is not obvious whether the appropriate "rationality" test should incorporate health shocks or not. The appropriate test depends on whether such shocks can/are accurately anticipated.

1. Rational Longevity Expectations when Cancer is Anticipated: What if health "shocks" are anticipated, or more accurately, the probability of suffering a health shock is known? For example, assume that there is a 50% chance of getting cancer and cancer inevitably leads to not being alive at 75. Furthermore, if an individual does not get cancer, there is also a 50% chance of surviving. If individuals accurately incorporate both of these probabilities into their longevity expectations, then longevity expectations should be 25% (50% chance of getting cancer and not being alive times the 50% chance of not getting chance and not being alive.) If these probabilities are correct and there are 4 individuals, then the number that are actually alive at 75 should be 1. So the data should look like the following.

Alive	LE	Error	Cancer
0	0.25	-0.25	1
0	0.25	-0.25	1
0	0.25	-0.25	0
1	0.25	0.75	0

Notice that when everyone correctly anticipates *both* the chance of getting cancer and the chance of being alive at 75, the average forecast error is 0, such that regressing Alive-LE on a constant results in the constant being 0, correctly suggesting rationality. Furthermore, regressing Alive-LE on a constant and a cancer "shock" would give a regression equation: Alive-LE = 0.25 – 0.5*Cancer. The constant is positive suggesting a pessimism bias, but this is *incorrect*, because cancer was rationally incorporated into longevity expectations and should not be included as a separate variable in the regression.

2. Rational Longevity Expectations when Cancer is Unanticipated: What if health shocks are unanticipated, or more accurately, the probability of suffering a health shock is completely unknown (and assumed to be 0%)? Assume that there is a 50% chance of getting cancer and cancer inevitably leads to not being alive at 75. Furthermore, if an individual does not get cancer, there is also a 50% chance of surviving. Finally, assume that individuals do NOT incorporate the probability of getting cancer because it is unknown. Then longevity expectations should be 50% (corresponding only to the 50% chance of not being alive.) If these probabilities are correct and there are 4 individuals, then the number that are actually alive at 75 should be 1. So the data should look like the following.

Alive	LE	Error	Cancer
0	0.5	-0.5	1
0	0.5	-0.5	1
0	0.5	-0.5	0
1	0.5	0.5	0

Notice that when cancer is unanticipated/unknown, the average forecast error is -0.25. However, it is possible that these individuals are still rational because the possibility of cancer was unknown when longevity expectations were made, so regressing Alive-LE on a constant results in the constant being -0.25, *incorrectly* suggesting pessimism. However, regressing Alive-LE on a constant and cancer would give a regression equation: Alive-LE = 0 – 0.5*Cancer. The constant is now zero, correctly suggesting rationality.

Appendix 2C: Numerical Example of Rationality Tests – Incorporating Known Information/Possible Sources of Bias

The purpose of the following example is to illustrate how the incorporation of information that is "known" at the time of the longevity expectation, but could possibly be a reason *for* the bias could be interpreted.

Optimistic Smokers Assume that non-smokers have a 50% chance of being alive at 75. Also, assume that 50% of individuals smoke and that smokers are guaranteed to not be alive at 75. However, further suppose that smokers ignore this knowledge and assume that they have the same chances of being alive at 75 as non-smokers. If these probabilities are correct and there are 4 individuals, then the number that are actually alive at 75 should be 1 (the two smokers are definitely not alive and of the remaining 2 non-smokers, 1 should be alive at 75.) So the data should look like the following.

Alive	LE	Error	Smoke
0	0.5	-0.5	1
0	0.5	-0.5	1
0	0.5	-0.5	0
1	0.5	0.5	0

Notice that the average error term is -0.25, suggesting aggregate optimism bias, such that a regression of Alive-LE on a constant results in a constant of -0.25. This correctly corresponds to the optimism that is associated with smokers.

Furthermore, the results of a regression of Alive-LE on both a constant and Smoke would be Alive-LE = 0 – 0.5*Smoke. Notice, the negative coefficient on Smoke (-0.5) corresponds to the optimism bias that is associated with smokers (who knew when forming their longevity expectations that they were smokers.) Also, note that when incorporating Smoke in the regression, the constant becomes 0. However, it is known that there is an aggregate optimism bias. Thus, in regressions that include known information, which are potentially sources of bias in forming longevity expectations, the constant should NOT be considered a measure of aggregate bias anymore.

Note, in this example, if smokers correctly incorporate the information associated with smoking, the data "should" be as follows.

Alive	LE	Error	Smoke
0	0	0	1
0	0	0	1
0	0.5	-0.5	0
1	0.5	0.5	0

Notice, the average error term here is again 0, suggesting rationality.