The Rationality of Retirement Expectations and the Role of New Information†

Hugo Benítez-Silva‡
SUNY-Stony Brook

and

Debra S. Dwyer
SUNY–Stony Brook

December 1, 2004

Abstract

This paper tests the rationality of retirement expectations, controlling for sample selection and reporting biases. We find that retirement expectations in the Health and Retirement Study are consistent with the Rational Expectations hypothesis. We also analyze how new information affects the evolution of retirement expectations and discover that, on average, individuals correctly anticipate most uncertain events when planning their retirement, except for some health shocks, the need for additional private health coverage, and the probability of a job change. Our results support a wide variety of models in economics that assume rational behavior.

Keywords: Retirement Expectations, New Information, Rational Expectations
JEL classification: D84, J26

† The authors are grateful to the Michigan Retirement Research Center (MRRC) and the TIAA-CREF Institute for financial support. A longer more detailed version of this paper is available in our websites or from the authors upon request. We are grateful to Moshe Buchinksy, John Rust, Tom Muench, Jim Hosek, Michael Hurd, Arie Kapteyn, Stan Panis, James P. Smith, Christian Belzil, Bryan Campbell, Nikolay Gospodinov, Jenny Hunt, Peter Arcidiacono, Ahmed Khwaja, Marjorie McElroy, Julia Lynn Coronado, Don Andrews, Valentina Corradi, Albert Marcet, Ernesto Villanueva, Douglas Fore, John Ameriks, Chris Swann, Olivia Mitchell, Caroline Austin, Anna Vellvé-Torras, an anonymous referee, and the editor for insightful comments and suggestions. We also would like to acknowledge excellent research assistance from Huan Ni, Selcuk Eren, and Wayne-Roy Gayle. We have benefited from comments from participants at the Society of Labor Economists Meetings in Baltimore in May 2002, the North American Meetings of the Econometric Society at UCLA in June 2002, the RAND Labor and Population group, the CIREQ Labor and Applied Econometrics Workshop at Concordia University, the Applied Microeconomics Seminar at Duke University, the Labor and Public Economics Seminar at UPF, and the Resident Scholar Program at the Board of Governors of the Federal Reserve System in Washington D.C. All remaining errors are the authors'.
‡ Department of Economics. SUNY-Stony Brook. Stony Brook, NY 11794-4384. Phone: (631) 632-7551, Fax: (631) 632-7516. hugo.benitez-silva@sunysb.edu and dsdwyer@notes.cc.sunysb.edu

1. Introduction

There is a growing body of research using expectations variables. In particular, retirement expectations are used for a variety of research endeavors, including the analysis of the connection between Social Security expectations and retirement savings (Lusardi 1999, Dominitz, Manski, and Heinz 2002), the relationship between retirement expectations and retirement outcomes (Bernheim 1989, Dwyer and Hu 1999, and Coronado and Perozek 2004), and the study of consumption patterns after retirement (Haider and Stephens 2003, and Hurd and Rohwedder 2003). On the other hand, the vast general retirement literature has had it as its main objective to understand the effect of social insurance programs on labor supply and wealth accumulation.[1] However, the dynamic nature of retirement expectations and the rationality assumption of those expectations—present in most models of retirement behavior at the micro level—have received relatively little attention. In this paper, we use the Health and Retirement Study (HRS) to directly analyze the evolution of retirement expectations and test whether they are consistent with the Rational Expectations hypothesis. We also examine the role of new information in retirement expectation formation by analyzing how well people anticipate changes over factors relevant to their retirement decisions.

There is a large literature that tests the rational expectations (RE) hypothesis using market level variables, but the evidence in favor or against the hypothesis is mixed. Below, we propose a slightly different approach in line with Bernheim (1990)'s model of expectations formation, to argue that testing whether past retirement expectations are a sufficient statistic for current retirement expectations is indeed a test of the rationality of those expectations. We find that our tests do not reject the rationality of retirement expectations. Therefore, we can continue to rely on rationality and the strategies used to model it, as a good first approximation to behavior by economic decision makers. When analyzing the role of new information we find that a model of perfect foresight is rejected with respect to some health and economic measures. Therefore, on average not all changes are anticipated.

The conceptual model and econometric specifications are presented in Section 2. Section 3 discusses the data used in the empirical analysis. Section 4 reports our main findings, and Section 5 concludes.

2. A Model of Retirement Expectations Formation, and a Test of Rationality

Suppose an individual and an econometrician are trying to predict a variable *X,* for example, retirement, which the individual has decided will be determined as a function of a sequence of random variables:

[1] For a review of the retirement literature see Lumsdaine and Mitchell (1999).

$$X = h(\omega_1, \omega_2, ..., \omega_T). \qquad (1)$$

The sequence of vector-valued variables inside the parenthesis is observed by the individual at time *t=1,2,…,T*. Then the individual will take action *X* after some or all the ω_t's have been observed. Let $\Omega_t = \{\omega_t\}_{t=1}^{t}$ be the information known at period *t* and let $\omega_t = \left(\omega_t^1, \omega_t^2\right)$, where both components of ω_t are observed by the individual, but only ω_t^1 is observed by the econometrician. Let then $\Omega_t^1 = \left\{\omega_t^1\right\}_{t=1}^{t}$. Then we define

$$X_t^e = E\left\langle X \middle| \Omega_t \right\rangle, \qquad (2)$$

where the variable *X*, retirement age in our case, is forecasted, the full information set at time *t* is represented by Ω_t, and *E* is the expectations operator. This guarantees that errors in expectations will be uncorrelated with the set of variables known at time *t*. Variables included in the vector representing the information set Ω, come from standard life cycle models of retirement behavior which establish that the factors that influence retirement (and we assume retirement expectations) include health, health insurance, and socio-economic status. Using the law of iterated expectations and assuming that the new information is correctly forecasted by agents, from equation (2) we get:

$$E\left\langle X_{t+1}^e \middle| \Omega_t \right\rangle = E[E\left\langle X \middle| \Omega_t, \omega_{t+1} \right\rangle | \Omega_t] = E\left\langle X \middle| \Omega_t \right\rangle = X_t^e, \qquad (3)$$

where ω_{t+1} represents information that becomes available between period *t* and *t+1*. Notice that it is essential here to assume that new information (its conditional distribution, not just its mean) is correctly forecasted. Without this additional assumption expression (3) would not be correct. We are going to test this assumption jointly with the RE hypothesis, once we also assume linearity of the process presented in (3). Notice that the assumption of correct forecasting is in essence no different from the assumption in the early RE literature that forecast errors are normally distributed with mean zero.

From (3), we can write the evolution of expectations through time as

$$X_{t+1}^e = X_t^e + \eta_{t+1}, \qquad (4)$$

where $\eta_{t+1} = X_{t+1}^e - E[X_{t+1}^e \mid \Omega_t]$, and therefore $E(\eta_{t+1}|\Omega_t)=0$. Notice that η_{t+1} is a function of the new information ω_{t+1} received since period *t*. Based on this characterization of the evolution of expectations, as in Bernheim (1990), we can test the rationality of retirement expectations with the following regression:

$$X_{t+1,i}^e = \alpha + \beta X_{t,i}^e + \gamma\, \Omega_{t,i}^1 + \varepsilon_{t+1,i}, \qquad (5)$$

where α is a constant, and γ is a vector of parameters that estimate the effect of information in period *t* on the next period's expectations. The RE hypothesis implies that α=γ=0, and β=1.[2] A weak RE test assumes that γ is equal to a vector of zeros, and tests for α=0 and β=1. In other words, it tests whether expectations follow a random walk. The strong RE test is less restrictive and also tests for γ=0.

We are also interested in the role of new information in the expectations formation process. We want to know whether new information affects the evolution of expectations, and whether this new information is to some degree anticipated. This connects with the growing literature on the role of uncertainty in how individuals plan their retirement. Even assuming we cannot reject the prediction of the RE hypothesis, namely that in (5), β=1, a characterization of the role of new information in the evolution of retirement expectations implies that individuals integrate old and new information with weights attached to each of these sources, when updating their expectations. Assuming the linearity of the function that individuals use to process information regarding the variables of interest, and further assuming that the elements of the information set Ω are jointly normally distributed, we can then write

$$X^e_{t+1,i} = X^e_{t,i} + \gamma\ (\omega_{t+1} - E[\omega_{t+1} \mid \Omega^1_t]), \qquad (6)$$

where the term in parenthesis is the difference between the observed new information and the expectation of this information individuals had in the previous period. If we had all these elements, we could then estimate (6) and would expect a coefficient of one in front of the retirement expectations as of time *t*. Also, the coefficients in γ (which we will assume are time invariant) could be interpreted as the effect of the unanticipated components of new information on changes in expectations. However, we do not observe the second element in the brackets. Although it could be estimated by making further assumptions about the relationship between new and current information, this is likely to be a very noisy procedure. Instead, we estimate the following equation

$$X^e_{t+1,i} = \alpha + \beta X^e_{t,i} + \gamma\, \omega_{t+1,i} + \varepsilon_{t+1,i}, \qquad (7)$$

where there is no reason to believe that β would be equal to 1: since the unobserved expectation term from (6) would enter into the error term, leading to possible biases in the coefficients of interest, which in part can be treated as a problem of measurement error and attenuated using IV techniques. Also, the estimates of the vector γ, are interpreted as the effects that changes in specific factors of Ω^1 have on expectations formation. A nonzero

[2] A value of α different from zero can be consistent with the RE hypothesis in a cross-sectional analysis in the presence of a macro shock. We try to control for this type of shocks using year dummies in our empirical analysis.

coefficient of a given change in a member of the information set can be interpreted as information not perfectly anticipated, and therefore not embedded in the individuals' expectations as of period *t*.

Econometric Specifications

Estimating equations (5) and (7) is in principle straightforward, but the likely presence of measurement error in the dependent variable and its lag, and sample selection complicate the methodology.[3] As with all survey data, measurement error in proxy variables is likely to exist. We will assume that the measurement error that individuals incur is in no way correlated with the rationality of their expectations formation process but has more to do with, for example, the different environments individuals faced in each wave of the panel. In the survey, individuals responded with a retirement expectation either as an expected retirement age, an expected year of retirement, or in years left to expected retirement. The last two ways of responding are likely to create considerable rounding errors, since they do not allow individuals to report a month or fractional answers. This component of the self-reports can bias the coefficients of interest in an unpredictable way. In order to eliminate this noise, we want to capture the true component of the expectation and purge it of this source of bias. If measurement error was not a problem, we would expect the β coefficient of the IV estimator to be very close to the one from OLS, assuming validity of the instrument set. However, we will see below that the coefficient significantly changes in the IV specification, and approximates the value predicted by the theory. Nothing constrains the β coefficient of the IV specification to move towards 1, and the fact that it does, can be interpreted as support of our estimation strategy. We correct the measurement error problem using instrumental variables. We use time *t* subjective survival probabilities to age 85, and an indicator of smoking behavior as instruments (exclusion restrictions correlated with the rate of time preference) of retirement expectations.

---

[3] The possible presence of focal points in the retirement expectations variable can give rise to non-normal regression errors, since the distribution of the dependent variable, and the main independent variable could be considered bimodal. In general the results of Conditional Moments estimation are fairly robust to this problem, especially if the sample size is fairly large. The most important properties of the linear estimators (that they are the best linear unbiased estimators and consistent, and that the variance estimator is unbiased and consistent, allowing us to use conventional tests), survive the non-normality of the errors.

In the case of the selection problem, we will be making the implicit assumption that those who did not respond to the question of interest would use the same process to analyze information if they were to actually respond, as those who actually answered the question. A version of this assumption underlies most research that uses Heckman's (1979) work.

We follow Wooldridge (2002, p.567) to consistently estimate the effect of previous expectations on current expectations in the presence of measurement error and sample selection. And from (5), we write:

$$X^e_{t+1,i} = \alpha_1 + \beta X^e_{t,i} + \gamma_1 Z_{t,1i} + \varepsilon_{t+1,1i}, \tag{8}$$

$$X^e_{t,i} = \alpha_2 + \lambda_1 Z_{t,1i} + \gamma_2 Z_{t,2i} + \varepsilon_{t,2i}, \tag{9}$$

$$Y_i = 1(\alpha_3 + \gamma_3 Z_{t,3i} + \varepsilon_{3i} > 0), \tag{10}$$

where we first estimate the selection equation (10) using a probit specification, where Y_i is equal to one if both the retirement expectation in period t and the retirement expectation in period $t+1$ are observed. This procedure allows for an arbitrary correlation structure of the disturbances in the three equations. Among the many possible explanations why the disturbances in (10) could be correlated with the disturbances in (9) and (8), the one that seems more appealing to us is that those individuals who are more likely to have thought about retirement seem to have been exposed to some events or information, which we fail to capture with the set of exogenous variables that we use to estimate (10). Those unobserved events or information makes them more likely to expect to retire earlier, on average, than the expected retirement age they reported in the previous period. If this is the reason for the correlation, estimating an uncorrected model will lead to severe downward bias in β, since we would be left with a sample of people more likely to expect to retire earlier than previously reported. We would also expect biases in the other coefficients, including the constant. We believe that other reasons, such as the possible correlation between thinking about whether and when to retire, should be captured by the inclusion of age in the selection equation.

Z_3 in equation (10) includes all the exogenous variables and any exclusion restriction of the selection equation with respect to the structural equation. The exclusion restrictions include indicators for whether the father and mother of the respondent reached retirement age, and the age of the respondent. When estimating this Corrected IV specification, we assume that age is a proxy for the information set, which only matters in terms of making you more or less likely to think about retirement, but does not directly affect the expected retirement age. The reason for this is twofold: First, as Vella (1998, p. 135) discusses, including a variable which has a fairly small range can lead to the apparent linearity of the Inverse Mills' ratio, which can result in weak identification of the selection model, inflated second step standard errors, and unreliable estimates of the coefficients of interest. Second, on economic grounds, we believe it is reasonable to assume that age matters for whether you have thought about retirement and the timing of that retirement. But, consistent with the assumption we are testing, age should not be included when estimating the structural equation. Relaxing this assumption does not change the qualitative results presented in the paper. We then consistently estimate (8) by performing a modified 2SLS procedure, where the first stage includes as regressors the exogenous variables used in (10), the Inverse Mills' ratio from the probit equation, and any additional instruments (exclusion restrictions), Z_2 in (9), the validity of which will be tested.

In order to assess the effect of new information on expectations formation we consistently estimate (7) using the same Corrected IV procedure, which in this case means we estimate the system presented in equations (8) to (10) but with equation (8) being substituted by equation (7). All the other properties of the estimator we have discussed apply to this specification.

3. Data

The HRS is a nationally representative longitudinal survey of 7,700 households headed by an individual aged 51 to 61 as of the first round of interviews. In this analysis we use the first five waves of data, covering the period from 1992 to 2000. We include observations for respondents that are working in any wave, or are non-employed (but searching for jobs), who report retirement plans. We exclude respondents who do not report retirement plans for two consecutive waves, which results in a sample of more than 11,000 respondents and around 24,000 person-period observations, once missing values in the main variables of interest are considered. In each wave, respondents are asked when they plan to fully or partially depart from the labor force. They are also asked if they thought much about retirement. These questions are not mutually exclusive, but most of the people who have not thought about retirement do not report an expected age. A number of individuals report they will never retire, although these respondents often change their minds and later report an age. In this paper, we assign an age of 77 for those who reported they will never retire (estimated longevity).

Table 1 provides an analysis of how these retirement expectations compare with a number of retirement measures. Expected retirement ages are distributed similarly to actual retirement ages with peaks at ages 62 and 65, as well as a peak for the bunching at 77 for those who never plan to retire. Over time, these expectations converge to between 62 and 65 with fewer people maintaining plans of retirement before age 62 or after age 65. Besides the apparent focal points that ages 62 and 65 play when people form expectations, it seems clear that retirement expectations are measuring retirement itself. Finally, the last row of the table shows the frequency of changes in retirement expectations, which is quite large among those reporting valid ages in consecutive interviews. However, the average change is under a year, with a standard deviation of around 5 years.

4. Empirical Results

Tests of Rationality of Retirement Expectations

Table 2 reports the weak and strong RE tests for the full sample. The data support the weak and strong RE hypotheses only in an augmented model that corrects for sample selection and measurement error in the reports of expected retirement ages, resulting in a selection corrected IV specification. Notice that we perform an F-test based on the null hypothesis that $\beta=1$ in equation (5), to test the RE hypothesis. We obtain a coefficient for β of 1.05 for the weaker test, which given the precision of the estimate, does not allow us to reject the hypothesis that expectations follow a random walk. For the strong test, we estimate the full model of equations (8) to (10). The β parameter is estimated to be equal to 0.9397, which clearly fails to reject rationality. This Corrected IV technique seems to circumvent one of the traditional drawbacks of instrumental variables estimation: the large increase in standard errors in the IV estimates. The standard errors of the β parameter are less than half that of the uncorrected IV estimator, although still larger than in the OLS estimation.

Our identification strategy hinges on the exclusion restrictions that we have made regarding the fact that the smoking behavior and the self-assessed probability of living to age 85 affect retirement expectations at time *t* but are not correlated with the disturbances in the main equation (8). The best we can do to convince the reader of the validity of the instruments is to test them following the suggestions in Bound, Jaeger, and Baker (1995), and Staiger and Stock (1997). We find that we have robust instruments, with large F statistics in the first stage of the IV procedure, several times larger, especially for the Corrected IV procedure, than the minimum value (around 10) suggested in Staiger and Stock (1997) as a good rule of thumb to check whether we are in the presence of weak instruments (first stage results are available from the authors upon request). Also, the model is overidentified, which allows us to test whether our instruments are exogenous with respect to the error term in the structural equation. In all cases we cannot reject the overidentifying restrictions. The reported results are the product of estimating the system of equations via GMM, which is robust against unknown forms of heterokedasticity.

The rationality of our expectations variable also predicts that in the strong test the information available at time *t* should not be significant after controlling for time *t* expectations when estimating (5). After controlling for sample selection and measurement error we find that most of these factors are no longer significant. The joint hypotheses that all the coefficients are equal to zero cannot be rejected at the 5% significance level, and the same is true for the estimates of the constant (this is also true for the weak test), which validates the other predictions of the RE hypothesis regarding retirement expectations, and further supports the rationality of retirement expectations.

The Role of New Information

In this part of our analysis we are interested in the effect of new information on expected retirement age. We examine the effects of changes in all of the relevant factors on plans by estimating equation (7) but correcting for sample selection bias and measurement error, following the same strategy as in equations (8) to (10). We hypothesize that there will be no significant effects if, on average, people are able to completely anticipate the new information. A significant coefficient on any covariate does not indicate surprise, since it is easy to write an example where even if individuals perfectly know the distribution over the uncertain event, once this happens, the estimated coefficient would be different from zero. A statistically significant coefficient can be interpreted to mean that the individual could have known the probabilities, but not what will happen.

Table 3 reports the results of estimating equation (7) by OLS, and also with corrections for selection bias and measurement error in order to attenuate the biases resulting from the fact that we do not have all the information of the structural model in (6). We find that the coefficient on the expected retirement age reported in the previous period is not significantly different from 1, suggesting that we have greatly reduced the biases, and increased the reliability of the estimated coefficients. The only factors that significantly alter retirement plans are related to some health conditions, the availability of private health insurance, and job transitions. Focusing on the IV specification (which is the preferred one given the lack of significance of the Inverse Mills' ratio in the corrected-IV specification), we observe that individuals who have been newly diagnosed with diabetes or high blood pressure expect to delay retirement. One explanation for postponing retirement with the onset of diabetes or high blood pressure may be an increase in consumption necessary to cover health care requirements, particularly if health insurance is tied to work. Most changes in economic status have little effect, except for those individuals who have changed jobs or have returned to work, who expect to retire later than before the transition, capturing the role of employment uncertainty. Also, those with new access to private health insurance expect to retire later, which seems to capture possible offset of losses of coverage on the job, given that we are controlling for changes in health status and employment transitions. We conclude from these results that researchers need to pay special attention to incorporating the uncertainty of health shocks, employment uncertainty, and the role of health insurance in models that assume rational expectations.

5. Conclusions

In this paper we test the Rational Expectations hypothesis in the formation of retirement expectations, and cannot reject the RE hypothesis after controlling for reporting errors and sample selection. This result supports a variety of models that use the rationality assumption. We also examine the role of new information in retirement expectations formation. We focus on how well people anticipate changes over factors relevant to the retirement decision. A model of perfect foresight is rejected with respect to some health and economic variables. Therefore, on average not all changes are anticipated. We find that some components of health that are associated with various health conditions, the role of employment uncertainty, and effects of health insurance, warrant extra attention in models that assume some type of rationality.

The results in this analysis are meant to foster further discussion and research on the issues surrounding the role of expectations in the economic modeling of retirement behavior, and expectations in general. The methodology we present can be easily applied in many other contexts with repeated observations of expectations variables at the micro level such as fertility expectations and educational attainment expectations.

References

Bernheim, B.D. (1989): "The timing of retirement: A comparison of expectations and realizations." In *The Economics of Aging*. Ed. D. Wise. Chicago: University of Chicago Press, 335—355.

Bernheim, B.D. (1990): "How Do the Elderly Form Expectations? An Analysis of Responses to New Information." In *Issues in the Economics of Aging*. Ed. D. Wise. University of Chicago Press, 259—285.

Bound, J., D.A. Jaeger, and R.M. Baker (1995): "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, **90-430** 443—450.

Coronado, J.L, and M. Perozek (2004): "Wealth Effects and the Consumption of Leisure: Retirement Decisions During the Stock Market Boom of the 1990s," manuscript, Federal Reserve Board.

Dominitz, J., C.F. Manski, and J. Heinz (2002): "Social Security Expectations and Retirement Savings Decisions," NBER Working Paper **8718**.

Dwyer, D.S., and J. Hu. (1999): "The Relationship Between Retirement Expectations and Realizations: The Role of Health Shocks," in *Forecasting Retirement Needs and Retirement Wealth*. Olivia Mitchell, P. Brett Hammond, and Anna M. Rappaport (eds.). University of Pennsylvania Press.

Haider, S., and M. Stephens (2003): "Is there a Retirement-Consumption Puzzle? Evidence Using Subjective Retirement Expectations," manuscript, Michigan State University.

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

Hurd, M. and S. Rohwedder (2003): "The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending and Retirement," NBER Working Paper **9586**.

Lumsdaine, R.L., and O.S. Mitchell (1999): "New Developments in the Economic Analysis of Retirement," *Handbook of Labor Economics*, **3C** 3261—3307.

Lusardi, A (1999): "Information, Expectations, and Savings for Retirement", in *Behavioral Dimensions of Retirement Economics*, by Henry J Aaron Brookings Institute.

Staiger, D., and J.H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, **65-3** 557—586.

Vella, F. (1998): "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, **33-1** 127—169.

Wooldridge, J.M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Table 1. Distribution of Retirement Expectations and Actual Retirement.

	Exp. W1	Exp. W2	Exp. W3	Exp. W4	Exp. W5	Self-R Ret.	Social Sec.	Fully Retired	Full and Partial Retirement
Age <50	0.19%	0.25%	0.14%	-	0.23%	9.11%	0.7%	11.8%	13.1%
Age 50	0.54%	0.34%	0.39%	0.22%	0.11%	2.1%	0.2%	2.51%	2.73%
Age 51-54	1.76%	1.14%	0.97%	0.9%	0.69%	11.3%	1.13%	12.7%	14.1%
Age 55	6.07%	4.7%	2.64%	2.24%	1.25%	6.4%	1%	6.18%	6.54%
Age 56-59	7.09%	7.34%	7.19%	5.43%	3.37%	**20.5%**	5.34%	**21.2%**	**21.5%**
Age 60	8.71%	9.71%	7.73%	7.13%	4.68%	7.4%	4.91%	7.22%	6.81%
Age 61	1.24%	1.1%	1.53%	1.48%	1.31%	7.51%	5.57%	6.56%	6.37%
Age 62	**30.8%**	**29.9%**	**27.5%**	**27.3%**	**22.6%**	**15.8%**	**49.4%**	**13.5%**	**12.9%**
Age 63-64	3.62%	4.79%	4.66%	5.29%	5.99%	9.63%	16.4%	9.18%	7.84%
Age 65	**19.2%**	**20.4%**	**23.1%**	**23%**	**23.3%**	**5.67%**	**10.6%**	**4.77%**	**4.54%**
Age >65	20.8%	20.3%	24.3%	27%	35.4%	4.54%	4.63%	4.36%	3.45%
Never	17.1%	15.8%	17.1%	14.8%	17.3%	-	-	-	-
# Obs.	3,708	3,256	2,808	2,230	1,753	5,346	3,967	4,902	6,404
% Changing	-	54%	54.2%	52.6%	59.8%	-	-	-	-

W1 to W5: the expected retirement age reported in the respective rounds of data.
Self-R Ret.: reported retirement age when individuals are asked about their employment status.
Social Sec.: age at which they started to receive Social Security benefits.
Fully Retired: answer to a direct question regarding when did they fully withdrawn from the labor force.
Full and Partial Retirement: includes partial retirement in the above definition.
% Changing: Percentage changing their responses among those reporting valid retirement expectations in both waves.

Table 2. Tests of Rational Expectations

Variables	Pooled OLS	IV	Corrected IV
Weak RE Test (H_0: β=1):	***Reject***	***Reject***	***Cannot Reject***
Constant	31.112(1.170)**	20.388(11.655)*	-2.529(2.569)
Expected Retirement Age$_t$	0.520(0.018)**	0.687(0.183)**	1.050036(0.042123)**
Inverse Mills' Ratio	-	-	-0.293(0.4699)
Test of Over-Id Restrictions	-	Cannot Rej. P-v=.7710	Cannot Rej. P-v=.1830
Test of Weak Instruments	-	Reject P-v=.0000	Reject P-v=.0000
Strong RE Test (H_0: β =1):	***Reject***	***Reject***	***Cannot Reject***
Constant	20.725(1.347)**	10.345(6.655)	11.4678(7.4364)
Expected Retirement Age$_t$	0.390(0.021)**	0.673(0.170)	0.93978(0.08395)**
Inverse Mills' Ratio	-	-	-4.237(2.049)**
Economic factors at time t			
Net Worth (in $100,000)	0.003(0.016)	0.019(0.019)	0.029(0.018)
Respondent Income (in $1,000)	-0.001(0.001)	-0.001(0.001)	-0.003(0.001)**
No Health Insurance	0.918(0.396)**	0.505(0.584)	0.429(0.665)
Private Health Insurance	0.014(0.191)	-0.066(0.207)	-0.092(0.238)
Self-employed	0.869(0.277)**	0.434(0.323)	-0.082(0.306)
Pension	-0.821(0.182)**	-0.488(0.254)	-1.719(0.650)**
Financially Knowledgeable	0.012(0.169)	-0.066(0.171)	-0.242(0.189)
Health factors at time t			
Health limitation	0.108(0.200)	0.030(0.210)	-0.046(0.236)
Good-Very Good-Exc. Health	-0.364(0.253)	-0.231(0.255)	-0.499(0.306)
Doctor visits	-0.004(0.010)	0.001(0.010)	0.006(0.011)
High blood pressure	-0.115(0.178)	-0.106(0.189)	-0.190(0.222)
Diabetes problems	-0.497(0.291)*	-0.518(0.318)	-0.380(0.368)
Cancer	-1.552(0.533)**	-1.108(0.652)*	-0.562(0.770)
Stroke	-0.944(0.612)	-0.115(0.726)	0.609(0.766)
Heart Problems	0.002(0.291)	0.049(0.333)	0.082(0.384)
Arthritis	0.000(0.171)	0.021(0.184)	-0.029(0.211)
Difficulty walking multiple blocks	-0.477(0.287)*	-0.370(0.316)	-0.447(0.372)
Difficulty climbing stairs	0.282(0.375)	0.356(0.383)	0.259(0.431)
Demographic factors at time t			
Age	0.340(0.022)**	0.200(0.078)**	-
White	0.013(0.182)	-0.093(0.198)	-0.134(0.210)
Male	0.526(0.161)**	0.511(0.169)**	0.172(0.232)
Bachelor's Degree	0.425(0.195)**	0.376(0.189)**	0.305(0.203)
Professional Degree	-0.560(0.241)**	-0.466(0.224)**	-0.829(0.348)**
Married	-0.306(0.192)	-0.159(0.221)	-0.076(0.231)
Wave 1-2	0.1625(0.1701)	-0.0016(0.196)	-0.263(0.208)
Wave 2-3	0.1912(0.1866)	0.182(0.200)	-0.209(0.284)
Adj. R^2	0.328	-	-
Test of joint Significance of Covariates	Reject. P-v=.000	Reject P-v=.008	Cannot Rej. P-v=.0775
Test of Over-Id Restrictions	-	Cannot Rej. P-v=.5773	Cannot Rej. P-v=.5338
Test of Weak Instruments	-	Reject P-v=.0000	Reject P-v=.0000
Number of Observations	4,987	4,721	4,634

Table 3. New Information and Retirement Expectations

Variables	Pooled OLS	IV	Corrected IV
Expected Retirement Age$_t$	0.5074(0.020)**	1.011(0.066)**	1.077(0.0443)**
Inverse Mills' Ratio	-	-	-0.0657(0.557)
Constant	31.66(1.271)**	-0.487(4.18)	-4.629(2.655)*
Economic Factors			
Changes in Net Worth (in $100K)	-0.0099(0.016)	-0.0001(0.017)	-0.0009(0.017)
Changes in Earnings (in $1,000)	-0.0003(0.0006)	-0.0001(0.0007)	-0.00006(0.0007)
Changes in Priv. Health Insurance	0.192(0.183)	0.531(0.211)**	0.557(0.219)**
Changes in No Insurance Status	-0.498(0.429)	-0.532(0.518)	-0.464(0.552)
Changes in Pension	0.0035(0.324)	0.116(0.402)	0.109(0.421)
New Job Transition	1.541(.264)**	1.234(0.315)**	1.171(0.359)**
Health Factors			
Health limitation transitions[4]	-0.202(0.191)	-0.222(0.228)	-0.247(0.238)
Good-V.Gd-Exc.Hlt. transitions[5]	0.417(0.249)*	0.051(0.308)	0.0101(0.318)
Stroke transitions[6]	-0.679(0.867)	-1.315(0.979)	-1.532(1.015)
Heart problems transitions	0.112(0.294)	0.084(0.355)	0.036(0.377)
Cancer transitions	1.140(0.660)*	0.512(0.743)	0.487(0.787)
Diabetes transitions	0.983(0.358)**	0.764(0.444)*	0.719(0.476)
High blood pressure transitions	0.227(0.174)	0.361(0.212)*	0.378(0.22)*
Arthritis transitions	-0.129(0.151)	-0.089(0.185)	-0.087(0.193)
Doctor visits	-0.0005(0.009)	-0.009(0.011)	-0.011(0.012)
Difficulty climbing stairs	0.322(0.325)	0.119(0.396)	0.107(0.406)
Difficulty walking multiple blocks	-0.06180.308)	-0.4841(0.366)	-0.493(0.383)
Demographic Factors			
Marriage transitions[7]	-0.3937(1.021)	-0.7257(1.123)	-1.438(1.127)
Adj. R^2	0.2633	-	-
Test of Over-Id Restrictions	-	Cannot Rej. P-v=.293	Cannot Rej. P-v=.3368
Test of Weak Instruments	-	Reject P-v=.0000	Reject P-v=.0000
Number of Observations	4,755	4,422	4,340

[4] =1 if new limitation, 0 if no change, and -1 if you got better
[5] positive means health improved
[6] positive means condition (Stroke, high blood pressure) worsened or newly diagnosed (0 no change, -1 if better)
[7] =1 means new marriage, = 0 no change, =-1 dissolved marriage (widow, divorced)