Using Subjective Conditional Expectations to Estimate the Effect of Health on Retirement

Pamela Giustinelli
Bocconi University

Matthew D. Shapiro
University of Michigan

Prepared for the 19th Annual Joint Meeting of the Retirement Research Consortium
August 3-4, 2017
Washington, DC

The research reported herein was pursuant to a grant from the U.S. Social Security Administration (SSA), funded as part of the Retirement Research Consortium (RRC). The findings and conclusions expressed are solely those of the authors and do not represent the views of SSA, any agency of the federal government, Bocconi University, the University of Michigan, the Michigan Retirement Research Center at the University of Michigan, or the Vanguard Group Inc. The authors would like to thank Feiya Shao and Ann Rodgers for excellent research assistantship. The research uses data from the Vanguard Research Initiative (VRI) that was developed by a research team under a program project grant from the National Institute on Aging (P01-AG026571). Vanguard supported the data collection of the VRI. Vanguard's Client Insight Group and IPSOS SA were responsible for implementing the VRI survey and provided substantial input into its design. John Ameriks, Andrew Caplin, and Matthew D. Shapiro are co-principal investigators of the VRI. The VRI design benefited from the collaboration and assistance of Joseph Briggs, Wandi Bruine de Bruin, Alycia Chin, Mi Luo, Minjoon Lee, Brooke Helppie McFall, Ann Rodgers, and Christopher Tonetti as part of the program project and from Annette Bonner (Vanguard) and Wendy O'Connell (IPSOS SA). This project uses Survey 4 of the VRI, which was designed by Ameriks, Briggs, Caplin, Lee, Shapiro, and Tonetti. For documentation of the VRI, including a dynamic link to the survey instrument, see http://ebp-projects.isr.umich.edu/VRI/

Introduction

The future solvency of the U.S. Social Security program is threatened by projected costs exceeding revenues. Policies to encourage older American to work longer could make Social Security more sustainable. The feasibility and effectiveness of increasing retirement age hinges on workers' ability to work longer, which in turn depends crucially on how workers' health evolves as they age. Consider a worker in their mid-late fifties who is today at greater risk of entering work-preventing health states than his contemporaries in past cohorts (e.g., Case and Deaton (2015), Banerjee and Blau (2013); however, see Auerbach et al. (2017) for high-income). Will this worker still be working as he reaches 62 and beyond? Consider a 62 year old worker in excellent health. Will this worker continue to work into his late sixties or early seventies? A standard way to answer these questions is to use the health states and retirement ages realized in an earlier cohort of workers to recover the relationship between the two variables. The main challenge to this approach is that for each individual we only observe the age at which they retired given the health path they actually experienced. What would each worker have done, had they been on a different health path?

Our paper provides a novel strategy for assessing the effects of changing health. Older workers participating in the Vanguard Research Initiative (VRI) are asked to report the conditional likelihood (on a 0-100 percent chance scale) that they will be working to specified horizons under alternative health scenarios. They also report their unconditional likelihoods of working to those horizons and of experiencing those health states. Using these data our analysis delivers novel, individual and aggregate level, estimates of the subjective *ex ante* treatment effect (SATE) of health on retirement age, given by the difference between respondents' likelihoods of working in low versus high health. Such estimates are a key ingredient of any forecast of the population labor supply given a conjectured health distribution and, thus, are of great relevance to the Social Security Administration.

We find that the SATEs of health on labor supply at 2 and 4 years horizons equal 0 for almost 30 percent of the respondents. The remaining 70 percent reports subjective expectations which imply a strictly negative SATE (median = -40 percent and std. dev. = 24 percent for the 2-year horizon; median = -30 percent and std. dev. = 25 percent for the 4-year horizon). A rich set of covariates and the unconditional expectations measures shed light on dimensions of heterogeneity in STEs

Relationship with the Literature

The determinants of retirement have been widely studied in economics and elsewhere.[1] The role of health has been subject to much debate due to the difficulties of unpacking the "health-work nexus." First, the sign of the relationship is theoretically ambiguous. Health might operate through a variety of mechanisms (e.g., productivity, utility, horizon, beliefs about those). Health might take many forms (e.g., expected trajectory vs. unexpected shocks, earlier vs. later changes, types of conditions).[2] Second, the magnitude of the relationship is hard to quantify empirically as health and work are jointly determined and feed dynamically into each other.[3] Finally, both retirement and health are subject to several measurement issues exacerbating the challenges of studying their relationship.[4]

Most studies have sought to identify causal effects off health shocks, taking advantage of rich longitudinal information on retirement, health states, and other covariates in the Health and Retirement Study (HRS), the Survey of Health, Aging, and Retirement in Europe (SHARE), the English Longitudinal Study of Aging (ELSA), and similar large-scale surveys.[5] This approach, however, relies on observation of a sufficient number of shocks in the data and makes it challenging to extrapolate to individuals who have yet to experience any shock.

Fewer studies have tackled the problem using retirement expectations or hypothetical choices.[6] McGarry (2004) studies the effect of health on work expectations of working HRS respondents and finds large health effects. She focuses on workers to avoid "justification bias" in self-reported health among retired individuals.[7] Kapteyn et al. (2007) and van Soest and

---

[1] Coile (2015) and Fisher et al. (2016) provide recent reviews.

[2] See Grossman (1972), Bound et al. (1999), Lumsdaine and Mitchell (1999), Blundell et al. (2016), and references therein.

[3] Indeed, a recent and growing set of works has focused on the effect of retirement on health; e.g., see Rohwedder and Willis (2010), Coe and Zamarro (2011), Behncke (2012).

[4] See Bound (1991), Dwyer and Mitchell (1999), Lindeboom and Kerkhofs (2009), and Kapteyn and Meijer (2014) on health measurement issues and Gustman et al. (1995, 2010) and Maestas (2010) on concepts and measures of retirement.

[5] This strategy has been used in both structural models and more reduced-form analyses; e.g., see Bound et al. (1999), Disney et al. (2006), McGeary (2009), van der Klaauw and Wolpin (2010), Garcia-Gomez (2011), Blundell et al. (2016), and Jones et al. (2016). See also Shao (2016) for construction of expectation-based shocks.

[6] Stated preference analysis (SP), whereby individuals are asked to rank or rate choice alternatives in multiple counterfactual choice sets, has a long tradition in marketing, transportation, and other fields, but has gained acceptance only recently in economics. Ben-Akiva and Lerman (1985), Louviere et al. (2000), Train (2003), Hensher et al. (2005), and others provide textbook treatments of SP methods and combined RP-SP methods. See also Barsky et al. (1997), Blass et al. (2010), and Ameriks et al. (2015b) for economic applications.

[7] McGarry (2004) investigates the role of contemporaneous and lagged health, health changes, objective conditions, and survival expectations on the subjective probability of working past age 62 and changes in this measure.

Vonkova (2014) study preferences for full and partial retirement in the Netherlands using hypothetical choices.[8] These latter papers recognize that uncertainty about future health, employment, and other factors might play a role when respondents make their choice evaluations, nonetheless they do not incorporate these dimensions in the choice scenarios, nor they allow respondents to express uncertainty in the form of choice probabilities.[9]

Outline

We analyze a sample of healthy, older workers from the Vanguard Research Initiative (VRI) (described below), who are in the best position to work long(er) and, thus, should be of particular interest to the Social Security Administration. Like McGarry (2004), we exploit the subjective work probabilities of working respondents; however, in addition to the unconditional expectations we designed and collected new measures of work expectations contingent on alternative health scenarios as well as subjective health expectations. Using our new measures, we address the following questions:

1. Will these workers have the health capacity to work longer?
2. Will these workers work longer?
3. What is the causal effect of health on work?

To answer 1, we analyze respondents' expectations about their health 2 and 4 years ahead and the heterogeneity of health expectations across individuals' characteristics. Specific features of the distributions of respondents' health expectations (e.g., the mean) can be interpreted as health forecasts for the relevant population of currently working individuals (e.g., the proportion of current workers who will be in high vs. low health in 2 or 4 years).[10] Thus, these forecasts provide population-level estimates of current workers' capacity to work at specified future horizons.

To answer 2, we analyze respondents' expectations of working in 2 and 4 years and the heterogeneity of working expectations across individuals' characteristics. Once again, features of the distributions of respondents' working expectations provide population forecasts of the labor supply at specified future horizons. Notice that these estimates take the probability

[8] Maestas (2010) uses working expectations in the HRS to study transitions into retirement and the phenomenon of partial retirement.
[9] See Manski (1999) for a motivation and Blass et al. (2010) and Delavande and Manski (2015) for applications.
[10] The idea of using subjective expectations to forecast population behavior or outcomes dates back to Juster (1966).

distribution of future health into account, as respondents must “integrate” over the alternative health states they might face at the specified horizons in order to report their subjective *unconditional* probability of working at those horizons.

To answer 3, we analyze respondents’ expectations of working in 2 and 4 years were they turn out to be in high health or, alternatively, in low health. The mean of the distribution of subjective working expectations conditional on facing high health in 2 (4) years is an estimate of the *counterfactual* proportion of current workers who would work in 2 (4) years if all of them happened to be in high health in 2 (4) years. Similarly, the mean of the distribution of subjective working expectations conditional on facing low health in 2 (4) years is an estimate of the *counterfactual* proportion of current workers who would work in 2 (4) years if all of them happened to be in low health in 2 (4) years. These quantities are counterfactual because only a fraction of current workers will actually end up being in high – alternatively in low – health in 2 (4) years, as estimated by the mean of the distribution of subjective health expectations. The difference between these two counterfactual subjective probability distributions yields the distribution of subjective *ex ante* treatment effects (SATE) of health on work at the individual level. Such effects can be then “aggregated up” to estimate various treatment effects at the population level (e.g., on the treated, on the untreated, moments of the distribution, etc.).[11] We especially investigate the distribution of sign and magnitude of such effects. Moreover, we take advantage of the rich set of covariates in the VRI to study how the SATEs vary across individuals’ characteristics.

Finally, we use our SATE estimates to simulate a hypothetical change in a person’s likelihood of entering low health in 2 and 4 years on population-level labor supply forecasts at those horizons. Specifically, we halve the likelihood of being in bad health for each person. We find that these hypothetical changes in the chances of entering low health increase the estimates of the fractions of individuals predicted to work in 2 years by 2 percentage points and in 4 years by 3 percentage points.

[11] See Arcidiacono et al. (2014) and Wiswall and Zafar (2016) for applications of the approach in a different domain.

Research Methodology

Subjective *unconditional* probabilities of working past specified ages collected in the Health and Retirement Study (HRS) have proved to be remarkably predictive of actual labor force participation at the population level (e.g., Hurd (2009)). Calculations using HRS data further reveal that subjective unconditional labor supply probabilities *stratified* by observed health conditions are in turn predictive of the actual labor supply conditional on health (e.g., Hurd and Rohwedder (2014)).

Our research project takes the natural next step and directly analyzes newly-collected measures of subjective labor supply probabilities *conditional* on alternative future health states. Hence, our novel approach is to combine expectations of health with expectations of labor supply. This approach is fundamental to the aim of using *ex ante* measurement on an entire population to study future decision when only a subset of the population will face a shock such as bad health. We have implemented such questions in the Vanguard Research Initiative (VRI). In this paper we analyze these questions and examine their implications for Social Security.

The Vanguard Research Initiative

The VRI is a survey-administrative linked dataset on older wealthholders. Survey respondents are account holders at the Vanguard Group who are aged 55 and above, are web-survey eligible, and have at least $10,000 in financial assets at Vanguard. As of December 2015, four surveys were completed by a panel of about 3,000 respondents, with each survey focusing on a different aspect of retirement decision-making. Our analysis is based mainly on Survey 4 (Labor).[12] We select our sample from respondents who meet the following criteria (i) who have taken the first 4 surveys of the VRI; (ii) who were working at the time of Survey 4 and, thus, eligible to answer the labor supply and health expectations battery, which provides the key measures for our analysis;[13] (iv) who gave complete and consistent responses to the latter battery; and (iv) reported being in high health in Survey 4.[14] This sample amounts to 970 respondents aged 57 to 81, currently in high health and working. Sample size decreases to 839 respondents for the

[12] Survey 1 (Wealth), Survey 2 (Long-term Care), and Survey 3 (Transfers) will provide relevant covariates.
[13] Some of these individuals had already retired from their career job and were working in a bridge job at the time of the survey. These individuals, too, were asked the expectations questions just described with reference to their bridge job.
[14] Fewer than 3 percent of respondents reported being in low health (i.e., fair or poor). We exclude this small group because it is easier to analyze and explain behavior in response to one direction of health shocks.

analysis of expectations with a 4 years horizon, which applies to individuals who reported a positive probability of working in 2 years. Table A1 in Appendix A summarizes the steps involved in selection of our analytic samples for the 2- and 4-year ahead question batteries.

The Analytic Sample and Key Survey Measures

The VRI respondents tend to be wealthier, more educated, and healthier than average. However, conditional on the screening conditions used to select the sample, they are similar to those from the HRS and the Survey of Consumer Finances (SCF) (Ameriks et al., 2015a). Our main analytic sample is further selected as it is made of working respondents in good or better health. This sample therefore represents a subpopulation of particular interest for our analysis, as it is made of individuals who in principle have the capacity to work longer but for whom assessing the casual link between health and retirement is particularly challenging as they have not experienced the health shocks used for identification in the standard approach.

Table 1 summarizes the main characteristics of the two analytic samples which we use for the 2 and 4 years analyses.

Table 1. *Sample Characteristics*

	2-Year Ahead	**4-Year Ahead**
Characteristic	**Percent**	**Percent**
Respondent's age (at VRI Survey 4)		
≤ 59	22.9	24.4
60-61	14	14.3
62	6.7	7
63-64	13.4	13.3
65	4.9	4.2
66-67	8.5	8.7
68-69	8.6	8.2
70-71	5.2	5.2
≥ 72	15.9	14.5
Respondent's gender		
Female	62.99	63.17
Male	37.01	36.83
Respondent's race/ethnicity		
Non-Hispanic white	94.74	94.87
Asian	2.68	2.86
Other	2.58	2.26

Respondent's marital status (at VRI S4)		
Partnered (married & share financial future)	65.5	64.8
Not partnered	34.5	35.2
Respondent's educational attainment		
High school or less	5.77	5.96
Some college	14.95	13.83
College graduate	38.97	38.38
Other advanced degree	19.59	20.50
MBA	7.94	8.46
JD, PhD, MD	12.78	12.87
Respondent's health status (at VRI S4)		
High (excellent, very good, or good)	100	100
Respondent's employment status (at VRI S4)		
Working (full-time or part-time)	100	100
Respondent's job type (at VRI S4)		
Career	60.62	61.50
Bridge	39.38	38.50
Respondent's occupation (at VRI S4)		
Management and professional	71.75	71.99
Other services	17.32	17.04
Operative	10.96	10.97
Total household wealth in USD (at VRI S4)		
First quintile	0 – 258,475	0 – 255,584
Second quintile	258,475 – 533,739	255,584 – 537,700
Third quintile	533,739 – 874,867	537,700 – 877,000
Fourth quintile	874,860 – 1,583,538	877,000 – 1,559,059
Fifth quintile	≥ 1,583,538	≥ 1,559,059
Replacement rate (Expected pension & SS; replacement rate, career job wage, at VRI S4)		
First quintile	0 – 24	0 – 24
Second quintile	24 – 39	24 – 39
Third quintile	39 – 58	39 – 58
Fourth quintile	58 – 87	58 – 88
Fifth quintile	87 +	88 +
Respondent's annual salary in USD (at VRI S4)		
First quintile	0 – 12,000	0 – 13,000
Second quintile	12,000 – 45,714	13,000 – 47,000
Third quintile	45,714 – 77,534	47,000 – 80,000
Fourth quintile	77,534 – 117,000	80,000 – 120,000
Fifth quintile	≥ 117,000	≥ 120,000
Spouse's age (at VRI S4)		
≤ 59	28.55	29.94
60-61	15.11	15.90
62	5.50	5.09

63-64	11.91	12.19
65	3.82	4.01
66-67	7.48	6.94
68-69	9.16	8.95
70-71	4.58	4.48
≥72	13.89	12.50
Sample size	635	544
Spouse's health status (at VRI S4)		
Excellent	21.54	21.45
Very good	44.40	44.75
Good	25.81	26.08
Fair	6.87	6.79
Poor	1.37	0.93
Sample size	635	544
Spouse's employment status (at VRI S4)		
Working (full-time or part-time)	48.85	49.07
Not working	51.15	50.93
Sample size	635	544
Sample size	970	839

Survey 4 begins by asking whether or not an individual is working. If so, it gets facts about the current job and establishes if it is the career job (Current job battery). If yes it gets information about whether the individual is searching for another job (On-the-job search battery). If not, it gets information about the career job, separation from it, and subsequent search (Career job, Separation, and Career-to-bridge search batteries). If not working, there is a similar sequence starting with information about last job. This sequence establishes information about career job, bridge job (if relevant), and the transitions and search.[15]

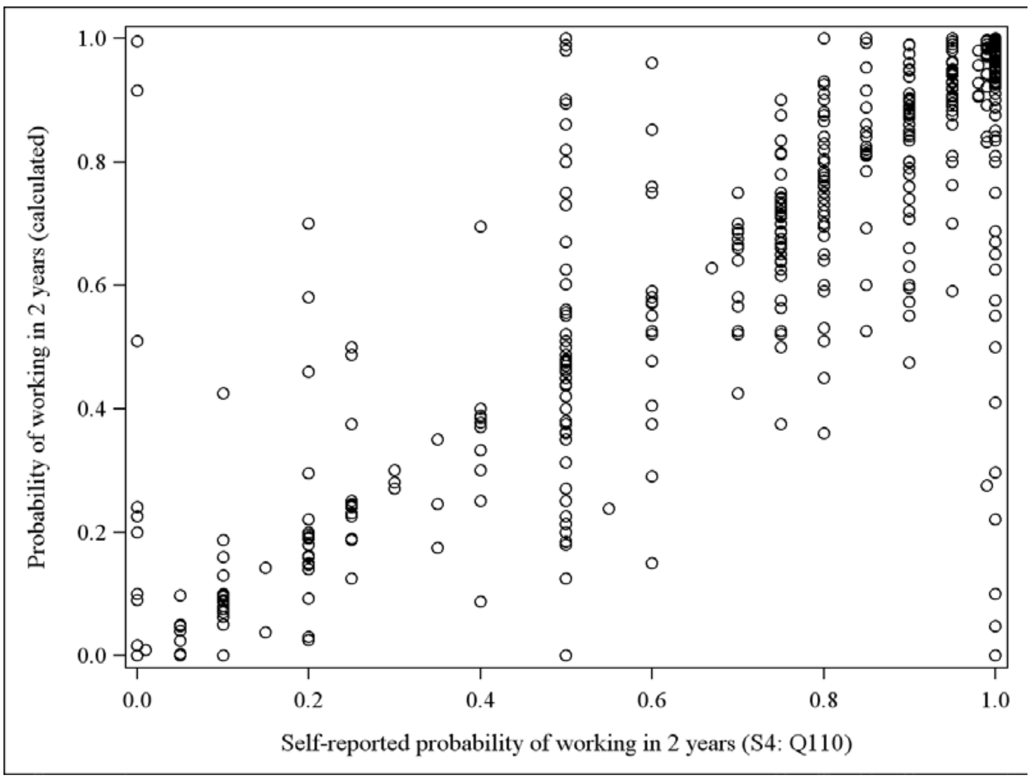
Respondents who were working in either a career job or bridge job at the time of Survey 4 were asked a series of questions regarding their labor supply and health expectations. Specifically, respondents were first asked for their self-rated health and, on a scale between 0 and 100, the percent chance that they will be working in 2 years from the point of the survey. Next, they were asked the percent chance that their health will be some particular state in 2 and 4 years. The partition of future health states used in the latter set of questions depends on the current level of health reported by the respondent (e.g., respondents in excellent health were

[15] In this paper we only use information about whether respondents worked in their career or bridge job at the time of survey 4 and about characteristics of their current job and, if different, their career job. In a separate paper we analyze job history, transitions, and reasons for transitions.

asked about their likelihood of being in very good/good and fair/poor health in 2 and 4 years). Finally, respondents were asked about their probability of working in the next 2 and 4 years, conditional on different health states.

Before analyzing each probability measure in details, we perform a preliminary analysis of the survey responses in order to provide immediate evidence supporting the validity of our novel measures. Specifically, we compare the unconditional labor supply probability given by each respondent with the value obtained by applying the Law of Total Probability (LTP) to the respondent's subjective conditional labor supply probabilities and the health probabilities. Figure 1 plots these two measures against each other.

Figure 1. *2-Year Working Expectations: Self-reported VS. LTP (Calculated)*



A large majority of the observations lies on the 45-degree line (or very close to it), corresponding to the case in which the self-reported probability and the calculated one are equal (or very close) to each other. The correlation between the two measures is approximately 0.95.Therefore, we conclude that overall respondents appear to understand the logic of probabilities and that we can dismiss potential concerns that the probabilistic reasoning needed to answer the questions is too difficult for our respondents.

As conditioning on health removes a potentially important source of uncertainty that respondents would otherwise need to factor in when answering the unconditional question, we hypothesize that subjective conditional probabilities might be even better predictors of conditional outcomes than unconditional probabilities are of the unconditional outcomes. A direct comparison between survey reports of the subjective conditional probability of working in 2-4 years with the conditional outcomes will become possible when future waves of the VRI are fielded and participants' labor supply observed.

Capacity to Work and Health Forecasts

Working VRI respondents are healthy, over 97 percent of them reports being in excellent, very good, or good health in Survey 4, which leaves us with a sample of 970 working individuals in high health (described in the first column of Table 1).

Do working VRI respondents expect to stay healthy or do they anticipate health declines? To answer this question we analyze respondents' expectations of being in high versus low health in 2 and 4 years elicited on a 0-100 scale of chance, where 0 means "no chance of being in high (low) health 2 (4) years from now" and 100 means "will be in high (low) health for sure 2 (4) years from now." Sample distributions (mean, standard deviation, and main quantiles) of percent-chance expectations are shown in Table 2 for all respondents in our analytic sample. Expectations are constructed from questions asking respondents to report the percent chance that they will be in a specific health state (or set of states) in 2 and 4 years, where the possible states are Excellent, Very Good, Good, Fair, and Poor and where Excellent, Very Good, and Good were subsequently consolidated into "high health" and Fair and Poor were consolidated in "low health."

Respondents' expectations of being in high health in 2 and 4 years range between 20 and 100 percent. Three fourths of the respondents report subjective likelihoods of being in high health in 2 years equal to or above 75 percent and of being in high health in 4 years equal to or above 70 percent. In fact, 10 percent of respondents expect to be in high health for sure in 2 years and 5 percent of them think that they will be in high health for sure in 4 years. Large fractions of healthy respondents express some uncertainty about their future health by reporting expectations strictly between 0 and 100 percent; about 90 percent for the 2 years horizon and 95 percent for the 4 years horizon.

Table 2. *2-Year and 4-Year Ahead Health Expectations*

	Percent chance of high health in 2 years	**Percent chance of low health in 2 years**	**Percent chance of high health in 4 years**	**Percent chance of low health in 4 years**
Mean	83.4	16.6	76.5	23.5
Std. Dev.	16.5	16.5	19.5	19.5
Q01	25	0	20	0
Q05	50	0	40	0
Q10	50	0	50	5
Q25	75	5	70	10
Median	90	10	80	20
Q75	95	25	90	30
Q90	100	50	95	50
Q95	100	50	100	60
Q99	100	75	100	90
N of obs.	970	970	839	839
% 0	0	13.61	0	6.56
% in (0, 100)	86.39	86.39	93.44	93.44
% 100	13.61	0	6.56	0
% 0	0	13.61	0	6.56
% in (0, 50)	1.75	75.15	6.67	74.37
% 50	9.48	9.48	12.40	12.40
% in (50, 100)	75.15	1.75	74.37	6.67
% 100	13.61	0	6.56	0

Next, we use box-and-whiskers plots to investigate whether these patterns vary by age, where age is displayed on the x-axis and the health probabilities on the y-axis. Figures 2A and 2B refer to the 2 years horizon and Figure 3A and 3B to the 4 years horizon. Figures 2A and 3A refer to the probability of high health and Figures 2B and 3B to the probability of low health. Each figure displays 9 box-and-whiskers plots, each corresponding to an age group.

With a 2-year horizon, the median health probabilities are remarkably constant across age groups.

With a 4 -year horizon, the median probabilities are still fairly stable among respondents in their fifties and possibly early sixties, but tend to decline with age thereafter. Moreover, the cross-sectional variance of responses tends to increase with age for the 4-year horizon.

Figure 3A. *2-Year Ahead Probability of High Health, by Age*

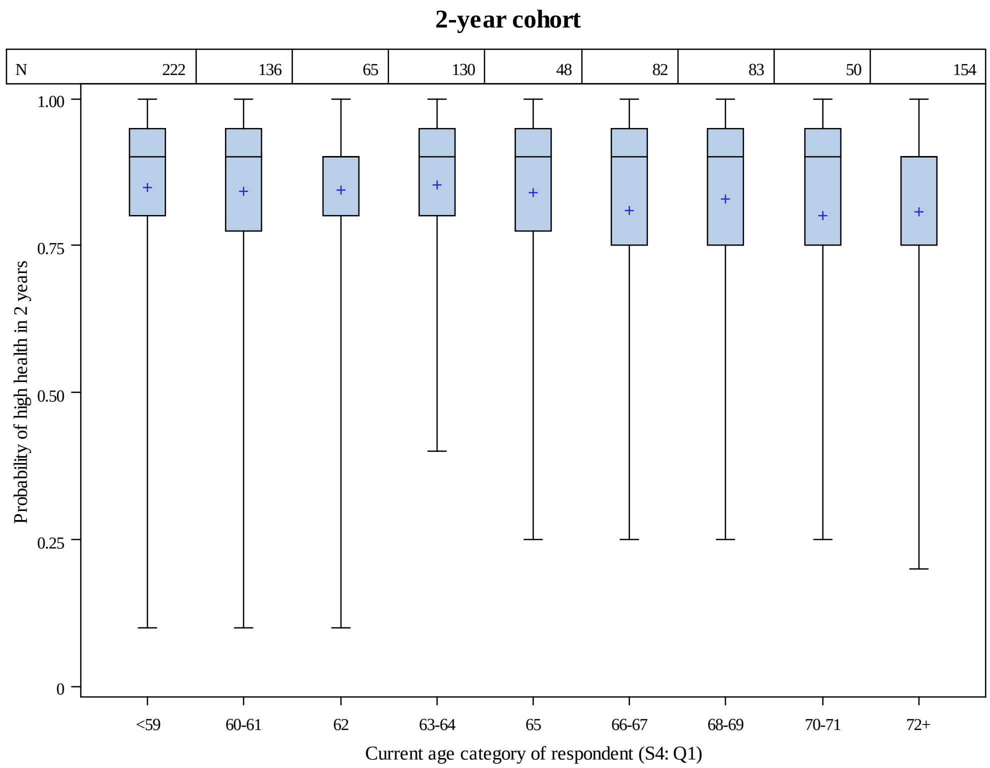


Figure 3B. *2-Year Ahead Probability of Low Health, by Age*

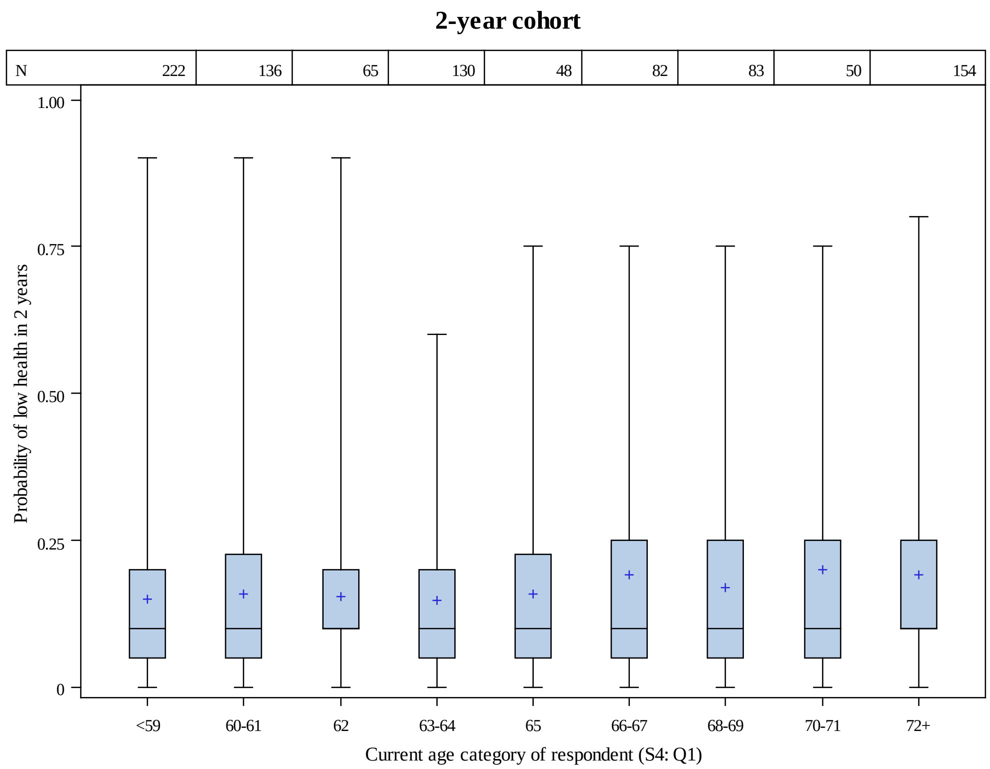


Figure 4A. *4-Year Ahead Probability of High Health, by Age*

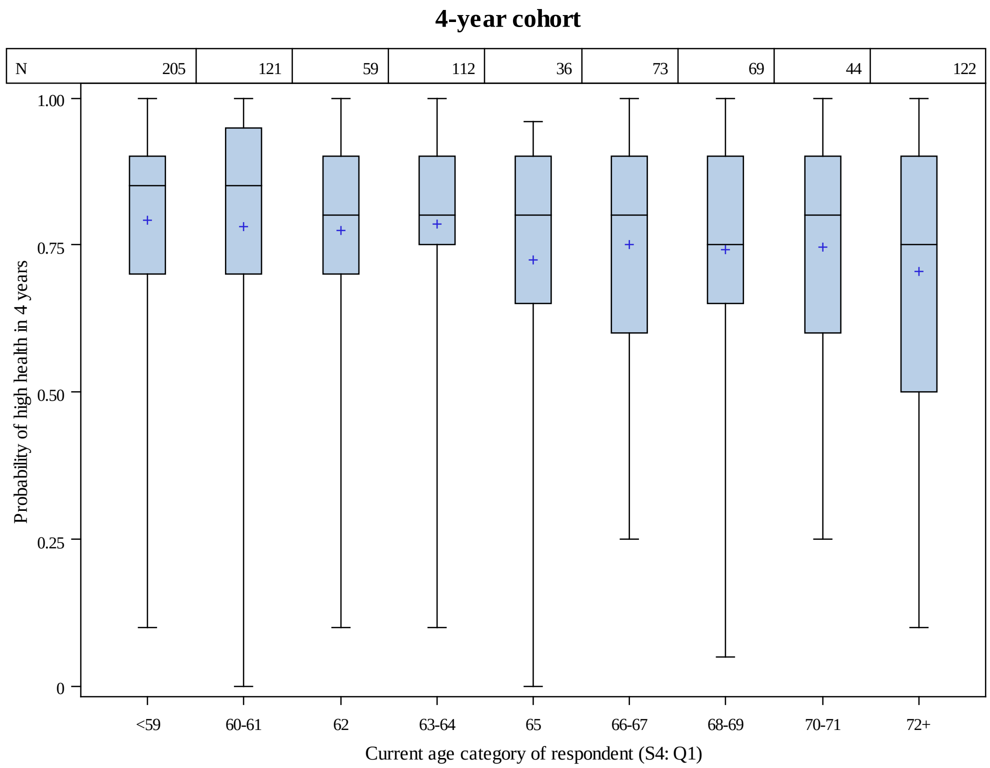
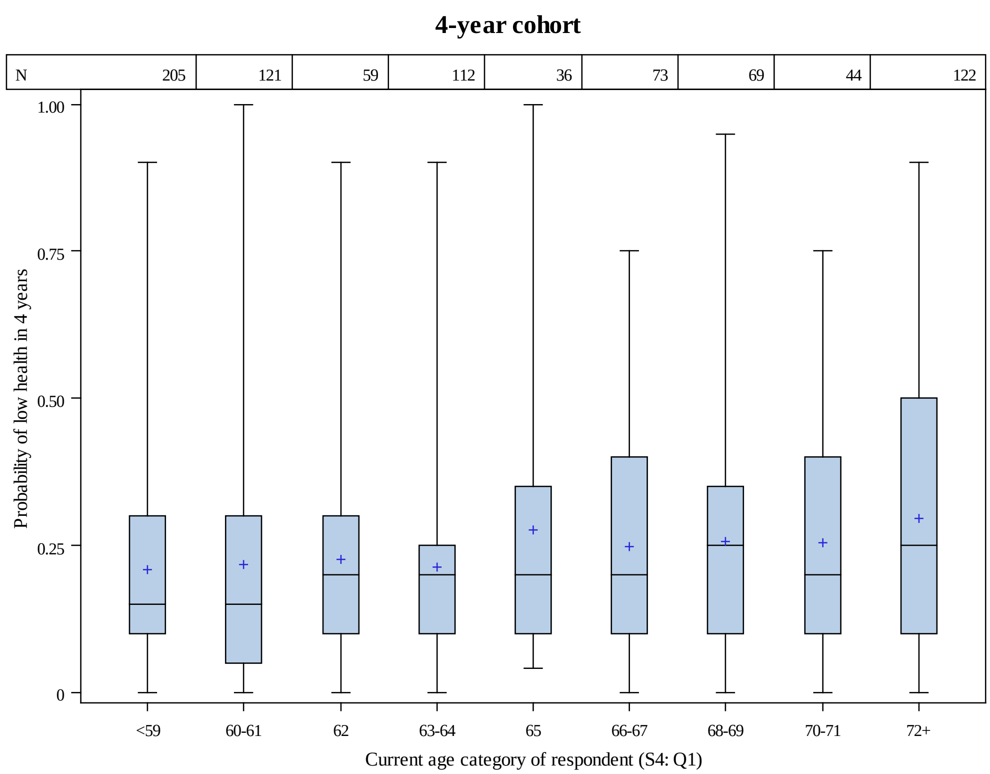


Figure 4B. *4-Year Ahead Probability of Low Health, by Age*



Health expectations might vary with other characteristics of the respondents in addition to their age. We investigate this possibility in Table 3. Specifically, in Table 3 we report best linear prediction estimates of the respondents' subjective probabilities of transitioning to low health in 2 and 4 years on the set of covariates described in Table 1. As all our predictors are dummy variables, estimated coefficients should be interpreted relative to the reference group. The latter is given by male respondents, aged 59 or younger, who have attained a high school diploma or a lower degree, currently working in their career job within the Management & Professional sector, who are not partnered, and who are in the highest quintile of the distributions of: total household wealth, current salary, and replacement rate.

Table 3. *Predictors of 2-Year and 4-Year Ahead Health Expectations*

	Probability of low health in 2 years			**Probability of low health in 4 years**		
Predictors	Coeff	SE	Signif	Coeff	SE	Signif
R's age						
age in 60-61	0.009	0.018		0.0007	0.022	
age = 62	0.003	0.024		0.018	0.029	
age in 63-64	-0.003	0.019		0.003	0.023	
age = 65	0.006	0.027		0.045	0.036	
age in 66-67	0.042	0.022	*	0.045	0.027	*
age in 68-69	0.014	0.022		0.054	0.028	*
age in 70-71	0.043	0.028		0.050	0.034	
age $\geq$ 72	0.039	0.020	*	0.097	0.026	***
R's gender						
female	-0.024	0.013	*	-0.042	0.016	***
R's education						
some college	-0.029	0.026		0.0008	0.033	
college grad	-0.028	0.025		-0.003	0.031	
other adv. degree	-0.036	0.027		-0.005	0.033	
MBA	-0.041	0.031		-0.0153	0.038	
JD, PhD, MD	-0.026	0.030		-0.007	0.037	
R's occupation						
operative	0.0003	0.015		0.007	0.019	
other services	0.039	0.019	**	0.069	0.024	***
R's job type						
bridge	-0.004	0.013		-0.017	0.016	
R's marital status						
partnered	-0.017	0.014		0.010	0.018	
Spouse's work status						
working	0.009	0.014		-0.004	0.017	

Total HH wealth						
1st quintile	0.010	0.020		0.036	0.025	
2nd quintile	-0.011	0.020		-0.006	0.024	
3rd quintile	0.014	0.018		0.014	0.023	
4th quintile	0.023	0.017		0.022	0.022	
R's replacement rate						
1st quintile	-0.008	0.019		0.005	0.023	
2nd quintile	-0.018	0.018		-0.002	0.023	
3rd quintile	-0.006	0.018		-0.001	0.023	
4th quintile	-0.018	0.018		-0.022	0.022	
R's current salary						
1st quintile	0.034	0.022		0.020	0.028	
2nd quintile	0.040	0.020	**	0.033	0.025	
3rd quintile	0.037	0.019	**	0.037	0.024	
4th quintile	0.021	0.018		0.026	0.022	
Constant	0.171	0.037	***	0.020	0.028	***
Sample size	970			839		
R^2	0.0416			0.0672		

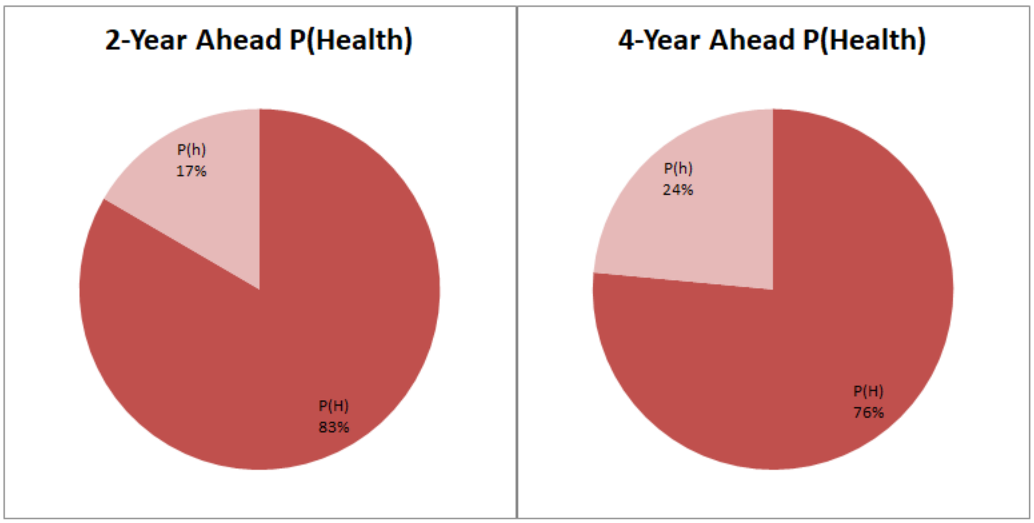
Note: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent.

The subjective probabilities of experiencing low health in 2 and 4 years tend to be higher among older respondents, respondents with lower-level occupations, and those with lower current salary and household wealth; whereas they tend to be lower among female respondents, respondents with higher educational qualifications, those who are partnered, those who are working in a bridge job, and respondents at lower quintiles of the replacement rate distribution. While the signs of these associations are generally intuitive, only the associations of the health probabilities with gender, occupation, and selected age categories and salary quintiles are statistically significant.

This limited observed heterogeneity of health expectations across respondents' characteristics can be easily rationalized in our context by recalling that our analytic samples are composed of relatively homogenous individuals all of whom were working at the time of the survey and in high health – arguably strong predictors of future health. They are also a highly relevant population for the working-longer question of this paper.

Overall, healthy working VRI respondents of all age groups appear fairly optimistic about their chances of remaining in high health – thus maintaining their capacity to work – as opposed to experiencing a health decline.[16] The means of the distributions of respondents' 2 years subjective probabilities of high and low health imply a forecast of the proportion of high health individuals, $P(H)$, equal to 83.4 percent and a forecast of the proportion of low health individuals, $P(h)$, equal to 16.6 percent. Similarly, the means of the 4 years distributions imply $P(H) = 76.5$ percent and $P(h) = 23.5$ percent respectively. These forecasts are shown in Figure 4 by means of pie charts.

Figure 4. *2- and 4- Year Ahead Health Forecasts, High (H) versus Low (h) Health*



Health forecasts at specified horizons are key ingredients for generating forecasts of labor supply at the corresponding horizons, to which we turn next.

Labor Supply Forecasts

Having established that a large majority of working VRI respondents expects to maintain high health – and thus the capacity to work longer – over the following 2 and even 4 years, it is natural to ask whether they also expect or plan to continue working. To answer this question, we analyze 2 and 4 years working expectations of these individuals elicited on a 0-100 scale of chance, where 0 means "no chance of working 2 (4) years from now" and 100 means "will work

[16] A direct comparison between expectations and realizations for these individuals will be possible when health realizations are observed in future VRI waves.

for sure 2 (4) years from now." Sample distributions (mean, standard deviation, and main quantiles) of percent-chance expectations are shown in Table 4.

Table 4. *2- and 4-Year Ahead Unconditional Working Expectations*

	Percent chance of working in 2 years (self-reported)	**Percent chance of working in 2 years (calculated)**	**Percent chance of working in 4 years (calculated)**
Mean	69.8	65.9	52.7
Std. Dev.	35.6	35.3	37
Q01	0	0	0
Q05	0	0	0
Q10	5	4.4	0
Q25	50	40	17
Median	90	80.1	50
Q75	100	97.5	90
Q90	100	100	100
Q95	100	100	100
Q99	100	100	100
N of obs.	970	970	839
% 0	9.48	9.18	11.68
% in (0, 100)	53.09	73.51	76.88
% 100	37.42	17.32	11.44
% 0	9.48	9.18	11.68
% in (0, 50)	12.16	21.86	36
% 50	11.96	2.16	2.5
% in (50, 100)	28.97	49.48	38.38
% 100	37.42	17.32	11.44

Respondents' working expectations at 2 and 4 years are somewhat heterogeneous and span the whole support of 0-100 percent chance scale. In column 1, over a third of the respondents (37.42 percent) expects that they will work for sure in 2 years, as opposed to the almost 10 percent of those who expect not to work for sure. The remaining majority of respondents (approx. 53 percent) views working in 2 years as an uncertain event. Nearly 12 percent of respondents report that they have 50 chances out of 100 of working in 2 years;[17] a

[17] Some of these respondents might be using 50 percent as an expression of "epistemic uncertainty" (e.g., see Fischhoff and Bruine de Bruin (1999) and Bruine de Bruin et al. (2002)), thereby conveying that they don't know or are unsure about their chances of working in two years. For this reason, the VRI has developed a response design for expectations questions that excludes 50 percent. For the module used in this paper, however, we use the 0-100 point scale for comparability with the HRS.

similar fraction gives a percent chance below 50 percent; whereas almost 30 percent gives a percent chance above 50 percent (the second most frequent response category after 100 percent). Indeed, the median belief of 90 percent is quite high, indicating that half of the respondents expect to work in 2 years with a likelihood of 90 percent or higher.

In column 2, the fraction of respondents who think that they will work for sure decreases to 17 percent, and the fraction of unsure respondents increases to 73 percent, when we use the "calculated" unconditional working expectations, (obtained by combining respondents' working expectations conditional on health and their health expectations according to the law of total probabilities), in place those reported directly by the respondents. However, the median probability is still very high and equal to 80 percent. Moreover, the medians, means, and standard deviations are very similar across the two distributions.

As the time horizon increases from 2 to 4 years in column 3, the fraction of respondents who think that they will not work for sure and the fraction of uncertain respondents increase respectively to 12 and 77 percent, whereas the fraction of respondents who think that they will work for sure decreases to 11 percent. While respondents' subjective probabilities of working in 4 years are still fairly high overall (e.g., the 75^{th} percentile is equal to 90 percent), these probabilities tend to be lower than those with a 2 years horizon (e.g., the median is down to 50 percent and the mean to 52.7 percent).

Do working expectations vary by age? To investigate this question, in Figures 5A-5C we create box-and-whiskers plots of the working expectations (on the y-axis) by current age of the respondent (on the x-axis). Age bins 60-61, 63-64, and 65 in Figures 5A and 5B are of particular interest, as a 2-year horizon from those ages implies the crossing of the early, normal, and full SS retirement ages (i.e., 62, 65, and 67), where actual labor supply displays well-known peaks. (Similarly, for age turns ≤ 59, 60-61 & 62, and 63-64 at the 4 years horizon in Figure 5C.)

In Figures 5A and 5B, the mean and median working expectations at 2 years feature sharp declines among the 60-61 years old (corresponding to the 62 peak), among the 63-64 years old (corresponding to the 65 peak), and among the 65 years old (corresponding to the 67 peak). Notice, however, that the mean and median working expectations do not decrease monotonically across groups of increasing age. This is consistent with increasing selectivity of the working and

Figure 5A. *2-Year Ahead Working Expectations, by Age, Self-Reported*

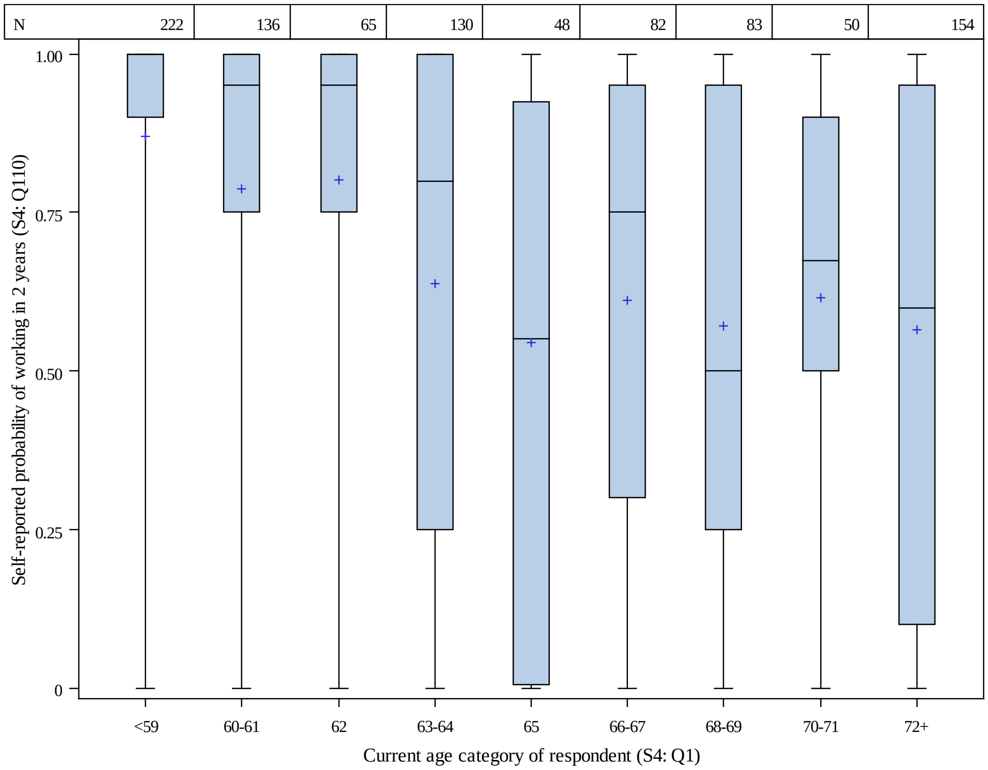


Figure 5B. *2-Year Ahead Working Expectations, by Age, LTC (Calculated)*

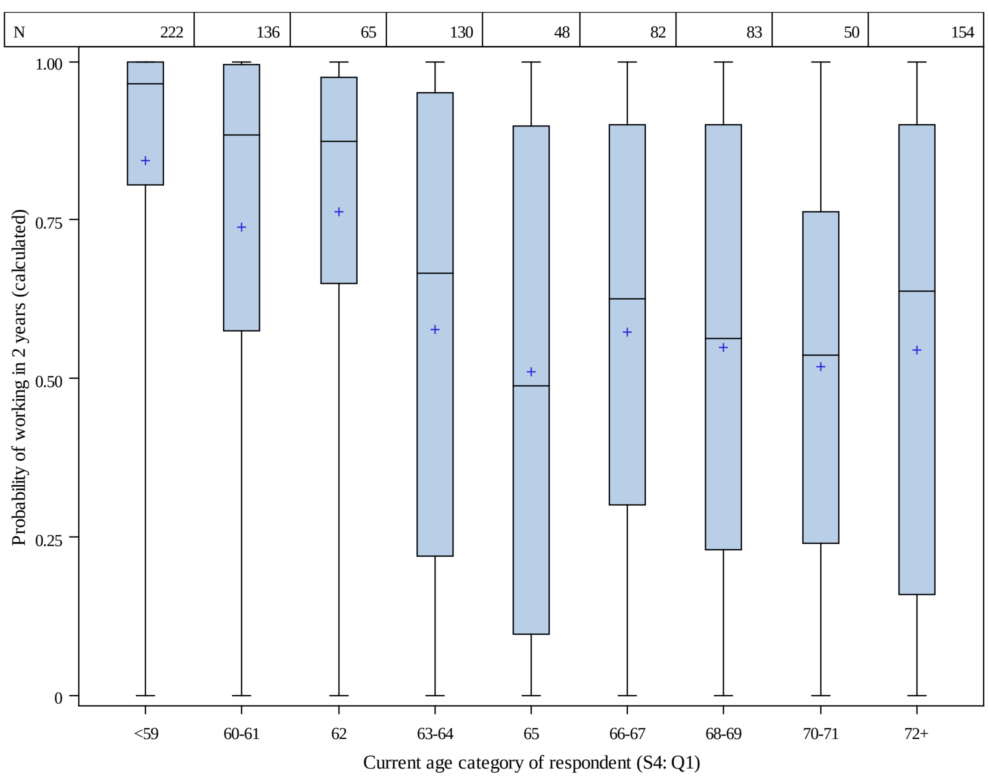
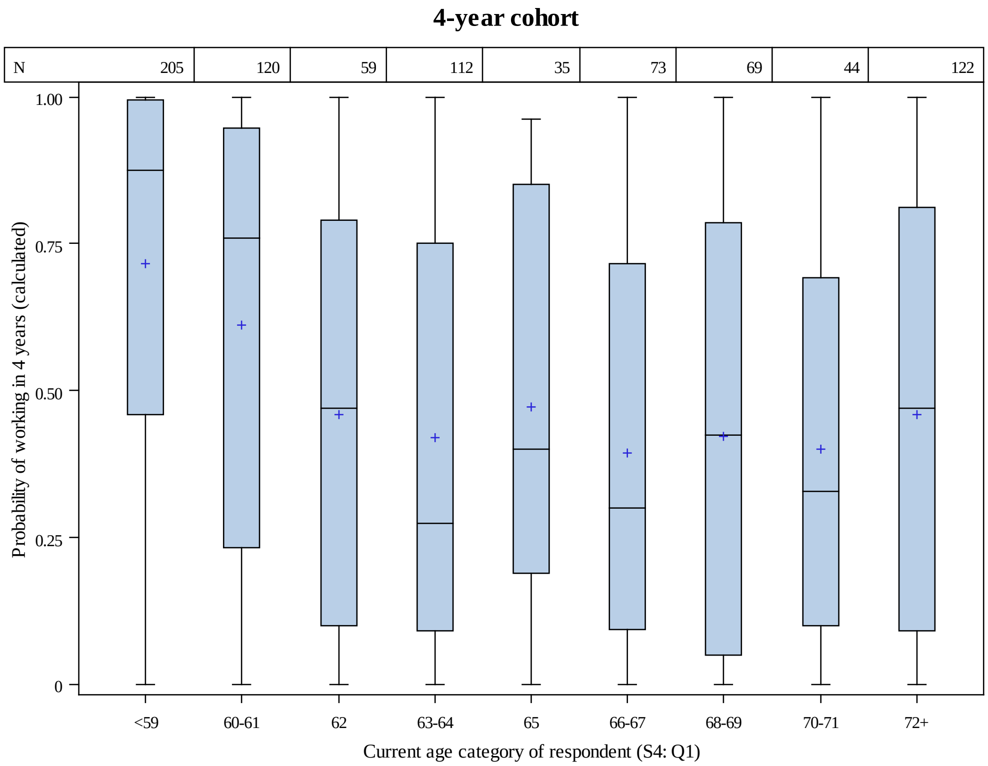


Figure 5C. *4-Year Ahead Working Expectations, by Age , LTC (Calculated)*



(high) health requirements applied to older respondents. An interesting additional feature of these age-specific distributions is that the cross-sectional variance of the subjective unconditional working expectations tends to increase with age, especially between respondents aged 62 or younger and respondents older than 62.

Moving to the 4-year-ahead horizon, inspection of Figure 5C reveals that the age-specific mean and median decrease sharply and steadily from the ≤ 59 and the 63-64 groups and level off (or tend to increase slightly) thereafter, again consistent with increasing selectivity of older sub-groups. The cross-sectional variance of working expectations is now fairly high in all age groups and appears higher than the cross-sectional variance of the 2 years working expectations. This is consistent with a higher uncertainty and/or a bigger role of heterogeneity as the forecasting horizon increases.

To ease the comparison between the 2 and 4 years, age-specific, unconditional working expectations distributions, Figures 6A and 6B combine them into the same graph in two alternative ways. Specifically, Figure 6A displays the distributions of 2- and 4-year ahead unconditional working probabilities by respondent's current age. Whereas Figure 6B displays the distributions of 2- and 4-year ahead unconditional working probabilities by the question's target age.

Figure 6A. *2- and 4-Year Ahead Working Expectations by Age Group*

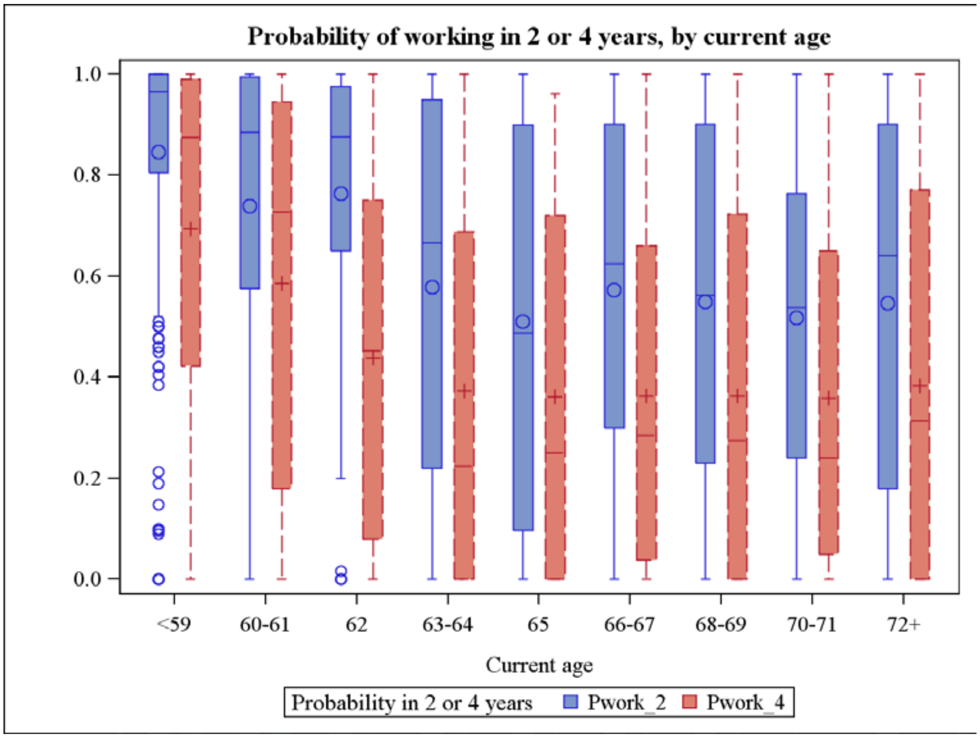
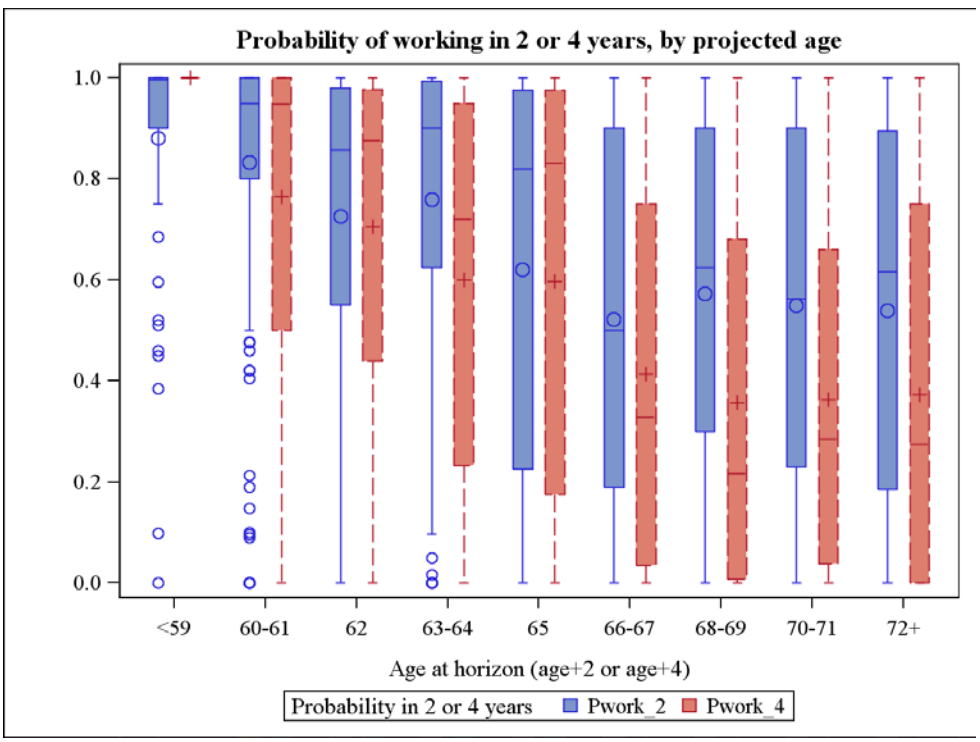


Figure 6B. *2- and 4-Year Ahead Working Expectations, by Projected Age*



Thus, while the 2- and 4-year ahead working expectations plotted in Figure 6A belong to the same respondents (within age category), in Figure 6B the 2 years probabilities belong to respondents who are 2 years older than the respondents whose 4 years probabilities are being compared with the latter (now within target or projected age).

Next we investigate whether 2- and 4-year ahead unconditional working expectations vary systematically by additional characteristics of our respondents beside age. As we did in Section 3 with the health expectations measures, we report parameter estimates of linear regressions of the unconditional working expectations on the usual set of covariates. Specifically, Table 5A shows estimated coefficients and standard errors from regressions of calculated unconditional working expectations at 2 years (column 1) and 4 years (column 2). Table 5B reports estimated coefficients and standard errors from a regression of self-reported unconditional working expectations at 2 years.

As in Table 3 above the reference group corresponds to male respondents, aged 59 or younger, who have attained a high school diploma or a lower degree, currently working in their career job within the Management & Professional sector, who are not partnered, and who are in the highest quintile of the distributions of: total household wealth, current salary, and replacement rate.

Table 5A. *Predictors of 2- and 4-Year Ahead Unconditional Working Expectations, Calculated*

Predictors	**Probability of working in 2 years (calculated)**			**Probability of working in 4 years (calculated)**		
	Coeff	SE	Signif	Coeff	SE	Signif
R's age						
age in 60-61	-0.089	0.036	**	-0.096	0.040	**
age = 62	-0.046	0.047		-0.232	0.052	***
age in 63-64	-0.239	0.036	***	-0.277	0.041	***
age = 65	-0.305	0.053	***	-0.239	0.064	***
age in 66-67	-0.242	0.043	***	-0.330	0.048	***
age in 68-69	-0.260	0.044	***	-0.302	0.050	***
age in 70-71	-0.264	0.054	***	-0.274	0.061	***
age $\geq$ 72	-0.228	0.040	***	-0.231	0.047	***
R's gender						
female	-0.011	0.025		0.014	0.028	

R's education						
some college	-0.018	0.052		0.033	0.059	
college grad	0.009	0.049		0.027	0.055	
other adv. degree	0.056	0.053		0.048	0.059	
MBA	0.048	0.060		0.045	0.067	
JD, PhD, MD	0.054	0.058		0.104	0.066	
R's occupation						
operative	0.012	0.030		-0.002	0.034	
other services	0.019	0.037		0.008	0.043	
R's job type						
bridge	-0.018	0.026		-0.0005	0.029	
R's marital status						
partnered	-0.067	0.028	**	-0.047	0.032	
Spouse's work status						
working	0.056	0.027	**	0.013	0.031	
Total HH wealth						
1st quintile	0.175	0.039	***	0.206	0.045	***
2nd quintile	0.137	0.037	***	0.207	0.043	***
3rd quintile	0.083	0.035	**	0.089	0.041	**
4th quintile	0.053	0.034		0.062	0.039	
R's replacement rate						
1st quintile	0.069	0.037	*	0.097	0.042	**
2nd quintile	0.077	0.036	**	0.033	0.041	
3rd quintile	0.053	0.035		0.018	0.040	
4th quintile	-0.007	0.035		0.032	0.040	
R's current salary						
1st quintile	-0.100	0.043	**	-0.026	0.049	
2nd quintile	0.036	0.040		0.054	0.045	
3rd quintile	-0.023	0.037		0.002	0.042	
4th quintile	-0.046	0.035		-0.036	0.039	
Constant	0.722	0.072	***	0.536	0.082	***
Sample size	970			839		
R^2	0.1883			0.1808		

Note: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent.

Table 5B. *Predictors of 2-Year Ahead Unconditional Working Expectations, Self-Reported*

Predictors	**Probability of working in 2 years (self reported)**		
	Coeff	SE	Signif
R's age			
age in 60-61	-0.067	0.036	*
age = 62	-0.035	0.047	
age in 63-64	-0.204	0.037	***
age = 65	-0.303	0.053	***
age in 66-67	-0.231	0.044	***
age in 68-69	-0.268	0.044	***
age in 70-71	-0.197	0.055	***
age $\geq$ 72	-0.234	0.041	***
R's gender			
female	-0.018	0.025	
R's education			
some college	-0.036	0.052	
college grad	-0.009	0.049	
other adv. degree	0.042	0.053	
MBA	0.044	0.061	
JD, PhD, MD	0.050	0.059	
R's occupation			
operative	0.014	0.030	
other services	0.033	0.038	
R's job type			
bridge	-0.012	0.026	
R's marital status			
partnered	0.064	0.027	***
Spouse's work status			
working	0.064	0.027	**
Total HH wealth			
1st quintile	0.195	0.039	***
2nd quintile	0.158	0.038	***
3rd quintile	0.097	0.036	***
4th quintile	0.089	0.035	**
R's replacement rate			
1st quintile	0.055	0.037	
2nd quintile	0.054	0.037	
3rd quintile	0.055	0.036	
4th quintile	-0.004	0.035	

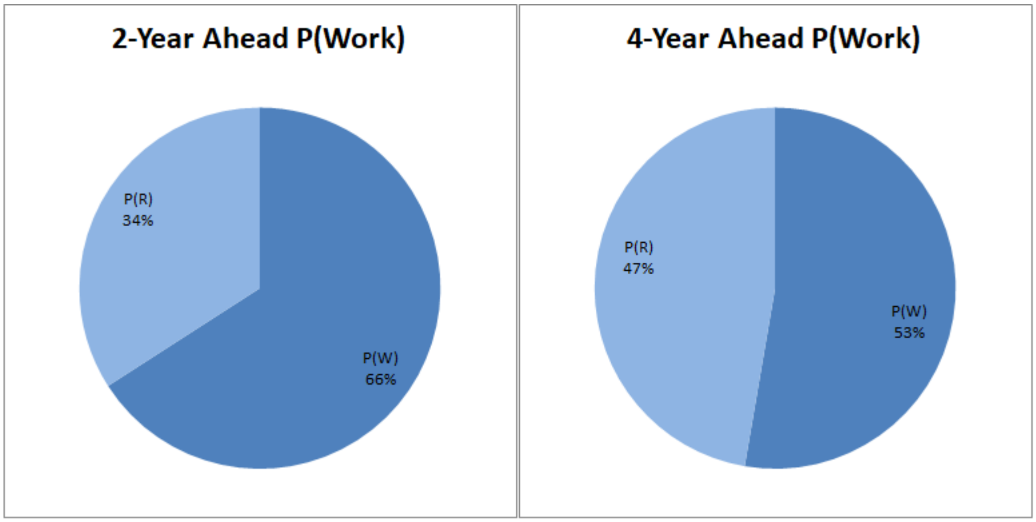
R's current salary			
1st quintile	-0.086	0.043	*
2nd quintile	0.052	0.040	
3rd quintile	-0.023	0.037	
4th quintile	-0.032	0.035	
Constant	0.749	0.073	***
Sample size	970		
R^2	0.1867		

Note: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent.

The subjective unconditional probabilities of working in 2 and 4 years are lower among older respondents, female respondents, partnered respondents, and respondents who are working in a bridge job. On the other hand, working expectations tend to be higher among more educated respondents, respondents working in lower-level sectors, respondents with a working spouse, and respondents at lower quintiles of the household wealth and replacement rate distributions. However, only the associations of the working probabilities with respondent's age, marital status, wealth, selected quintiles of the replacement rate, and spouse's working status are statistically significant.

Overall, the analysis of this section suggests that the working expectations or plans of healthy working individuals are somewhat sensitive to their age, family, and financial circumstances. In aggregate, however, the means of the distributions of respondents' subjective labor supply probabilities imply 2- and 4-year forecasts of the proportions of working individuals, $P(W)$, equal to 65-70 percent and 53 percent respectively. These forecasts are displayed by the pie charts of Figure 7.

Figure 7. *2- and 4-Year Ahead Labor Supply Forecasts*



The Causal Effect of Health on Work

The estimates shown in Figures 4 and 7 provide *status quo* population forecasts at 2 and 4 years of the health and labor supply of the kind of healthy working individuals represented by our sample. To be able to simulate alternative labor supply scenarios, represented by *hypothetical* or *counterfactual* labor supply distributions which would obtain under alternative health distributions, it is necessary to estimate the *causal effect* of health on work, that is, to measure the *structural* relationship between the two variables.

In this section, we first analyze respondents' subjective working expectations *conditional* on experiencing high and, alternatively, low health in 2 and 4 years. Then, we use these two measures of subjective conditional working expectations to derive subjective *ex ante* measures of the causal effect of health on work at both the individual and aggregate levels.

Working Forecasts If Everyone Is in High Health, If Everyone Is in Low Health

We elicited respondents' 2- and 4- year ahead working expectations *conditional* on alternative health scenarios by means of the familiar 0-100 scale of chance, where 0 means "no chance of working if in high (low) health 2 (4) years from now" and 100 means "will work for sure if in high (low) health 2 (4) years from now."

Table 6 displays the mean, standard deviation, and main quantiles of the sample distributions of percent-chance conditional working expectations for both health scenarios and prediction horizons.

Table 6. *2- and 4-Year Ahead Working Expectations Conditional on High and Low Health*

	Percent chance of working if in high health in 2 years	**Percent chance of working if in low health in 2 years**	**Percent chance of working if in high health in 4 years**	**Percent chance of working if in low health in 4 years**
Mean	70.5	41.9	58.7	33
Std. Dev.	36	36.1	39	34.4
Q01	0	0	0	0
Q05	0	0	0	0
Q10	5	0	0	0
Q25	50	5	20	0
Median	90	40	68	20
Q75	100	75	100	50
Q90	100	100	100	99
Q95	100	100	100	100
Q99	100	100	100	100
N of obs.	970	970	839	839
% 0	9.38	22.27	12.04	27.77
% in (0, 100)	50.10	64.33	56.73	62.34
% 100	40.52	13.40	31.23	9.89
% 0	9.38	22.27	12.04	27.77
% in (0, 50)	13.81	28.97	25.03	35.76
% 50	9.48	17.73	9.89	13.11
% in (50, 100)	26.80	17.63	21.81	13.47
% 100	40.52	13.40	31.23	9.89

In column 1, the distribution of working expectations conditional on experiencing high health in 2 years is remarkably similar to its unconditional counterpart. In particular, about 40 percent of respondents expect that they will work for sure in 2 years if their health is high; this proportion is only 3 percentage points higher than the 37 percent of respondents who report an unconditional working probability of 100 percent. The remaining respondents are split between a group close to 10 percent who gives a 0 percent chance of working in 2 years if in high health and the remaining 50 percent who gives a percent chance strictly between 0 and 100 percent. The first proportion is virtually identical to the proportion of 0s observed for the unconditional working probability. The second proportion is 3 percentage points lower than the corresponding proportion of responses to the unconditional question, mechanically compensating the slight increase in the 100 percent answers.[18] The mean of the distribution of conditional working expectations given high health is equal to 70.5, higher than both the mean of the distribution of unconditional working probabilities as directly reported by respondents (equal to 69.8 percent) and the mean of the distribution of calculated unconditional working probabilities (equal to 65.9 percent). The median working probability conditional on high health is equal to 90 percent and identical to the median probability of the distribution of self-reported unconditional working probabilities.

This close similarity between reports of unconditional working probabilities and conditional working probabilities given high health at 2 years is consistent with the observed high probability that the majority of respondents assigns to the event of experiencing high health in 2 years time. On the other hand, the distributions of the 4 years working probabilities, unconditional and conditional on high health, are less similar to each other than the 2 years distributions. This is also not surprising, as respondents' expectations of experiencing high health in 4 years are lower than the 2 years expectations, while featuring greater cross-sectional variation.

As the time horizon increases from 2 to 4 years between column 1 and column 3, the fraction of respondents who think that they will not work for sure and the fraction of uncertain respondents increase respectively to 12 and 57 percent, whereas the fraction of respondents who think that they will work for sure decreases to 31 percent. While the proportion of 0s is virtually identical to that observed for the distribution of the 4 years unconditional working probabilities in Table 4, the proportion of 100s is dramatically higher (31 percent versus 11 percent). In fact, both the mean and median probabilities of working in 4 years when experiencing high health (equal to 58.7 and 68 percent respectively) are visibly higher than the corresponding mean and median probabilities of the unconditional distribution (equal to 52.7 and 50 percent).

Having respondents entertain a scenario of low health – a negative and unexpected shock for the majority of them – lowers substantially their self-reported working expectations at both 2 years (in column 2) and 4 years (in column 4). For example, in the 2 years case the median of

[18] The fraction of 50 percent decreases slightly from 12 percent to 9 percent. A possible interpretation of this finding is that conditioning on future health reduces respondents' perceived uncertainty about their chances of working in the future.

the distribution of the conditional working probabilities drops from 90 to 40 percent between high and low health states. Similarly, the mean drops from 71 to 42 percent. In the 4 years case, the median drops from 68 to 20 percent and the mean from 59 to 33 percent.

Entertaining a low health scenario appears to additionally increase respondents' perceived uncertainty relative to the high health scenario; the fraction of respondents giving a response strictly between 0 and 100 percent increases from 50 percent to 64 percent under the 2 years horizon and from 58 percent to 62 percent under the 4 years horizon.

Overall, these figures suggest that, increasing older workers' chances of experiencing high health in the medium term (4 years) or, equivalently, reducing their chances of experiencing low health, might be an effective strategy for generating working-longer outcomes. In section 6 we explore this hypothesis more explicitly by simulating the effect on labor supply of a reduction in the respondents' 4 years subjective probabilities of experiencing low health by half.

The means of the distributions of the 2 and 4 years probabilities of working conditional on high health can be interpreted as 2 and 4 years labor supply forecasts if everyone were to experience high health at those horizons. The means of the distributions of the 2 and 4 years working probabilities conditional on low health have a symmetric interpretation. Figures 8A and 8B display these hypothetical forecasts.

Figure 8A. *2- and 4-Year Ahead Labor Supply Forecasts if Everyone Has High Health*

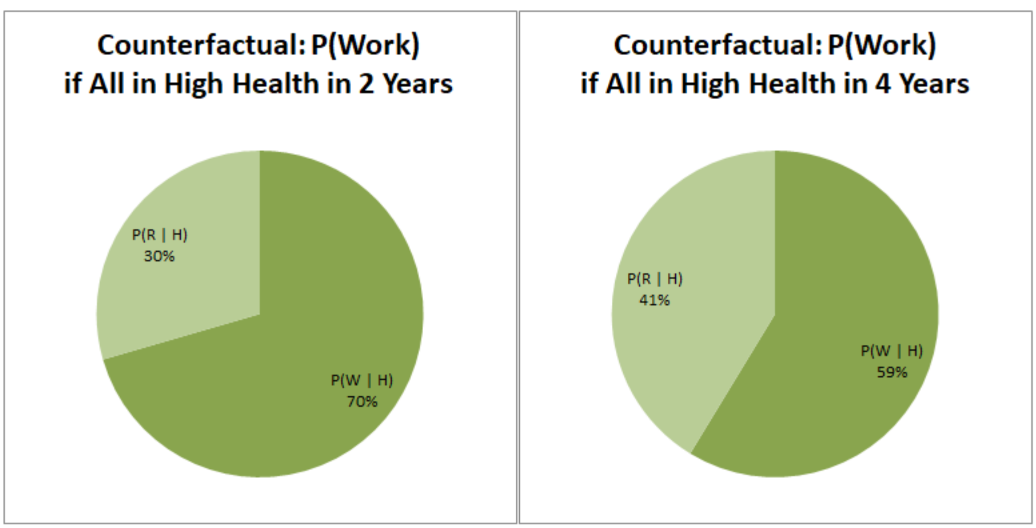
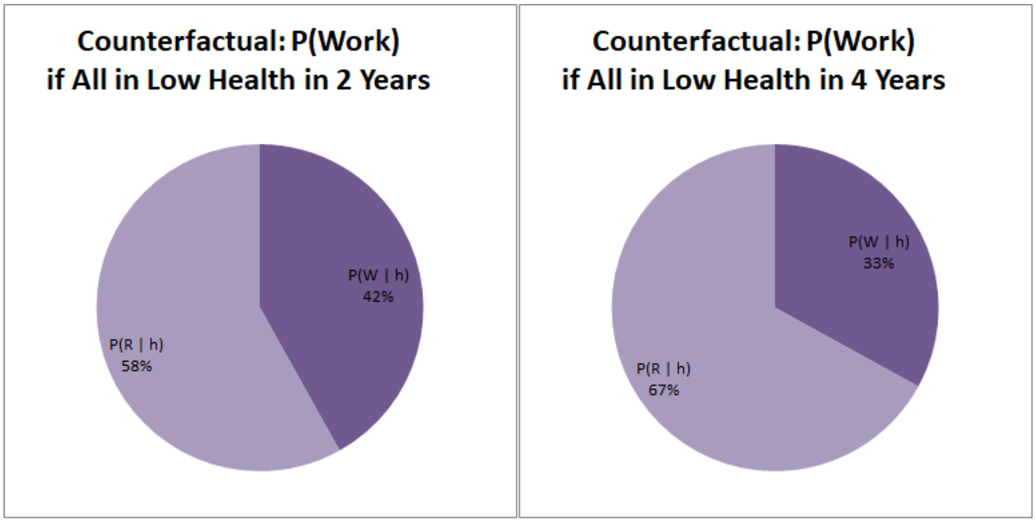


Figure 8B. *2- and 4-Year Ahead Labor Supply Forecasts if <u>Everyone Has Low Health</u>*



Subjective Ex Ante Treatment Effects of Health on Work

The subjective working probabilities conditional on experiencing low and, alternatively, high health are the two components of the individual-level subjective *ex ante* treatment effect of health on work (SATE). More explicitly, for each respondent i we measure their $SATE_i$ at a specified horizon,[19] as the difference between i's subjective probability of working if they were to experience low health at that horizon, $P_i(W \mid h)$, and i's subjective probability of working if they were to experience high health at that horizon, $P_i(W \mid H)$, that is,

$$SATE_i = P_i(W \mid h) - P_i(W \mid H).$$

Table 7 shows the sample distributions of the individual-level SATEs for the 2-year horizon (column 1) and the 4-year horizon (column 2).

Approximately 70 percent of the respondents gives a conditional subjective probability of working given low health which is strictly smaller than their subjective probability of working if in high health regardless of the horizon (i.e., SATE < 0). The vast majority of the remaining respondents (28-29 percent depending on the horizon) gives the same probability of working under the two health scenarios (implying a SATE = 0). Only a negligible fraction of respondents report a positive SATE. Note that a positive SATE is a theoretical possibility, e.g., if the value of leisure fall or need to work increases in bad health.

[19] The time index is omitted for simplicity.

Conditional on a negative SATE (columns 3 and 4), the distribution of the (absolute values of the) SATE ranges between 2-4 percent and 100 percent regardless of the horizon. The absolute values of both the median and mean SATE are somewhat higher in the 2-year horizon (40 and 41 percent) than in the 4-year horizon (30 and 37 percent), possibly because a negative health shock is a stronger bite at younger ages than at older ages, when additional motivations for stopping working might kick in.

Table 7. *2- and 4-Year Ahead Subjective Ex Ante Treatment Effects (SATE) of Health on Work*

	2-Year Ahead SATE	**4 Year-Ahead SATE**	**2 Year-Ahead SATE (if SATE < 0)**	**4-Year Ahead SATE (if SATE < 0)**
Mean	-28.5	-25.7	-40.9	-36.8
Std. Dev.	27.9	27.6	24.1	25.1
Q01	-100	-100	-100	-100
Q05	-80	-80	-90	-85
Q10	-70	-70	-75	-75
Q25	-50	-50	-50	-50
Median	-25	-20	-40	-30
Q75	0	0	-20	-15
Q90	0	0	-10	-10
Q95	0	0	-9.4	-5
Q99	5	0	-4.3	-2
N of obs.	970	839	682	594
% STE = -100	1.55	1.43		
% STE in (-100, 0)	68.76	69.37		
% STE = 0	28.45	28.25		
% STE > 0	1.24	0.95		
% STE = -100	1.55	1.43		
% STE in (-100, -	14.74	12.75		
50)	14.43	11.32		
% STE = 50	39.59	45.29		
% STE in (-50, 0)	28.45	28.25		
% STE = 0	1.24	0.95		
% STE > 0				

Note: SATE = probability of working if in low health – probability of working if in high health.

To investigate whether the main patterns observed in Table 7 vary by respondent's age, Figures 9A and 9B show box-and-whiskers plots of the 2- and 4-year ahead SATEs by age group. Overall, the mean and median SATE appear fairly stable across age groups, although a couple of dips are present at 62 and 63-64 among the 2-year ahead distributions and at 65 and 68-69 among the 4-year ahead distributions. The cross-sectional variance SATE is visibly wider at specific ages; 60-61 and 63-64 under the 2 years horizon and 60-61 and 66-67 under the 4-year horizon.

In Table 8 we study more systematically whether the individual-level SATEs for the 2- and 4-year horizon vary by respondents' characteristics. We report estimates from best linear predictions of 2- and 4-year ahead SATEs on the familiar list of covariates.

Figure 9A. *2-Year Ahead SATE, by Age*

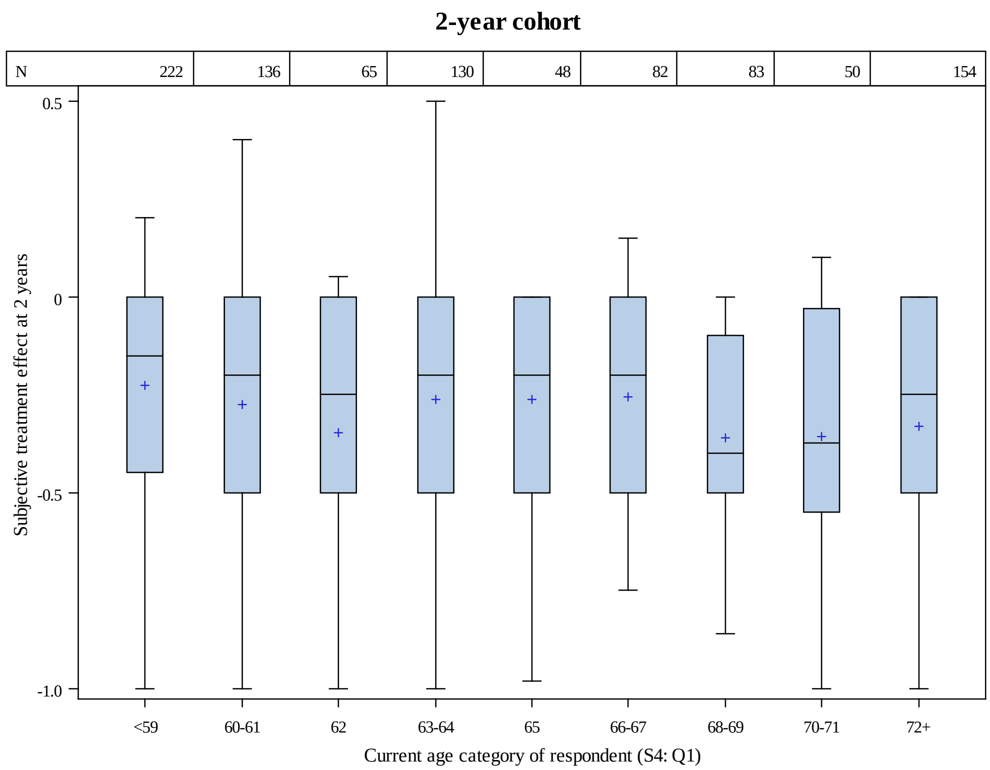
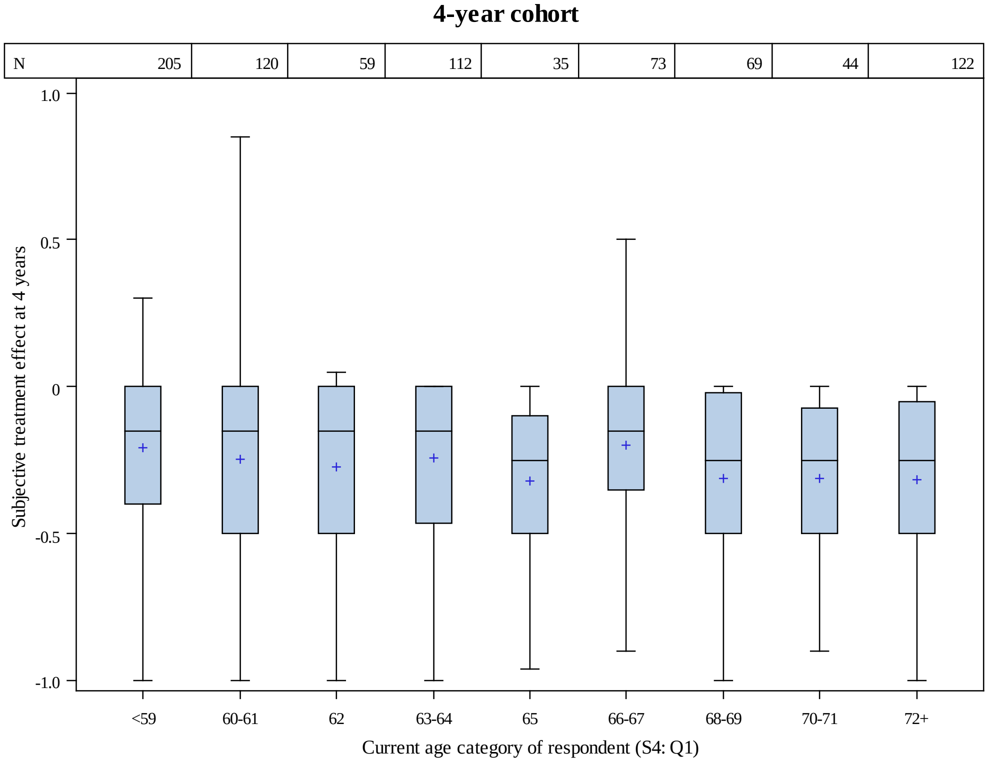


Figure 9B. *4-Year Ahead SATE, by Age*



At both horizons, the SATE tends to be larger (in absolute value) among older respondents, more educated respondents, partnered respondents, respondents with a working spouse, and respondents at lower quintiles of the household wealth, replacement rate, and salary distributions. The SATE tends to be smaller (in absolute value) among female respondents and respondents working in a bridge job or in lower-level sectors. These associations are statistically significant only for respondent's age and for selected quintiles of the replacement rate (in the 2-year case) and of the total household wealth (in the 4-year case).

Table 8. *Predictors of 2- and 4-Year Ahead SATEs*

	2-Year Ahead SATE			4-Year Ahead SATE		
Predictors	Coeff	SE	Signif	Coeff	SE	Signif
R's age						
age in 60-61	-0.046	0.031		-0.038	0.032	
age = 62	-0.117	0.040	***	-0.055	0.041	
age in 63-64	-0.034	0.031		-0.034	0.033	
age = 65	-0.031	0.045		-0.111	0.051	**
age in 66-67	-0.021	0.037		0.028	0.039	
age in 68-69	-0.120	0.037	***	-0.081	0.040	**
age in 70-71	-0.116	0.046	**	-0.088	0.049	*
age $\geq$ 72	-0.088	0.034	**	-0.086	0.037	**
R's gender						
female	0.001	0.021		-0.012	0.023	
R's education						
some college	-0.002	0.044		-0.025	0.047	
college grad	0.006	0.042		-0.010	0.044	
other adv. degree	-0.042	0.045		-0.019	0.047	
MBA	-0.014	0.051		0.003	0.054	
JD, PhD, MD	-0.031	0.050		-0.076	0.053	
R's occupation						
operative	0.008	0.025		-0.008	0.027	
other services	-0.020	0.032		-0.020	0.034	
R's job type						
bridge	0.008	0.022		-0.015	0.023	
R's marital status						
partnered	-0.012	0.024		-0.010	0.026	
Spouse's work status						
working	-0.014	0.023		-0.003	0.025	
Total HH wealth						
1st quintile	-0.039	0.033		-0.039	0.036	
2nd quintile	-0.045	0.032		-0.084	0.034	**
3rd quintile	-0.020	0.030		-0.032	0.032	
4th quintile	-0.044	0.029		-0.044	0.031	
R's replacement rate						
1st quintile	-0.013	0.031		-0.022	0.033	
2nd quintile	0.002	0.031		0.028	0.033	
3rd quintile	-0.023	0.030		-0.023	0.032	
4th quintile	-0.018	0.030		-0.018	0.032	
R's current salary						
1st quintile	-0.039	0.037		-0.032	0.039	
2nd quintile	-0.067	0.034	**	-0.040	0.036	
3rd quintile	-0.0001	0.032		0.008	0.034	
4th quintile	-0.005	0.030		-0.003	0.031	

Constant	-0.150	0.061	**	-0.116	0.065	*
Sample size	970			839		
R^2	0.0484			0.0528		

Note: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent.

Population Parameters

Aggregating up the individual-level SATEs yields subjective *ex ante* versions of population parameters well-known in the treatment-effect literature (e.g., see Heckman (2005)). Specifically, similar to Arcidiacono et al. (2014) and Wiswall and Zafar (2016), we compute estimates of the average subjective *ex ante* treatment effect (ASATE), the average subjective *ex ante* treatment effect on the treated (ASATT), and the average subjective *ex ante* treatment effect on the untreated (ASATU), as follows:

$$ASATE = \frac{\sum_{i=1}^{N} P_i\left(W \mid h\right) - P_i\left(W \mid H\right)}{N}$$

$$ASATT = \frac{\sum_{i=1}^{N} P_i\left(h\right)\left[P_i\left(W \mid h\right) - P_i\left(W \mid H\right)\right]}{\sum_{i=1}^{N} P_i\left(h\right)}$$

$$ASATU = \frac{\sum_{i=1}^{N}\left[1 - P_i\left(h\right)\right]\left[P_i\left(W \mid h\right) - P_i\left(W \mid H\right)\right]}{\sum_{i=1}^{N}\left[1 - P_i\left(h\right)\right]}$$

In the above expressions, $i = 1,...,N$ indexes the individuals. $P_i(h)$ denotes individual i's subjective probability of entering low health at the specified horizon (whose index is suppressed for simplicity). $P_i(W \mid h)$ and $P_i(W \mid H)$ denote individual i's subjective probability of working if they were to experience low and, respectively, high health at the specified horizon.

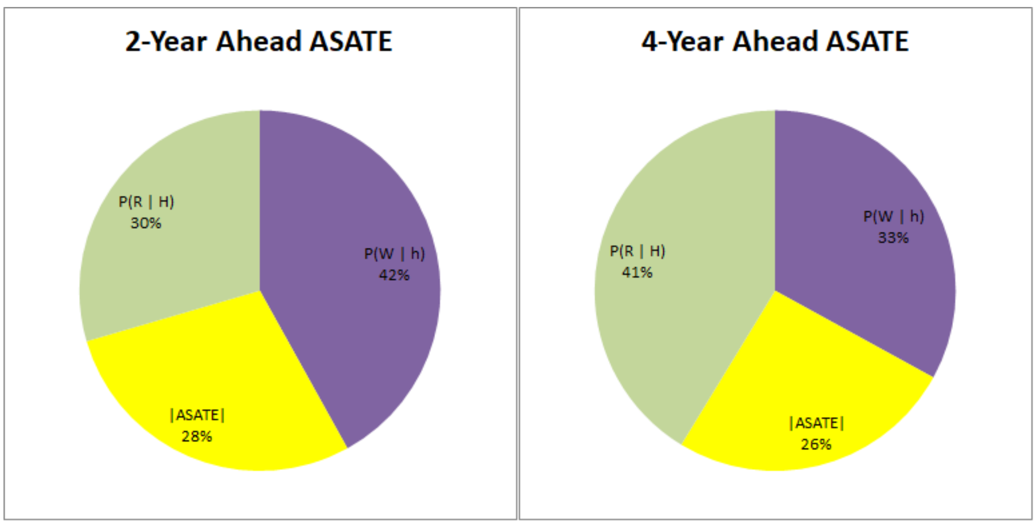
The calculated values of the three parameters for the two horizons are shown in Table 9.

Table 9. *2- and 4-Year Ahead Average Treatment Effect Parameters*

Horizon Aggregate TE parameter	**2 Years**	**4 Years**
ASATE	-28.53	-22.25
ASATT	-27.43	-25.16
ASATU	-28.75	-25.81

The ASATE can be visualized by putting Figures 8A and 8B on top of each other. Using the fact that for a fixed horizon, $|ASATE| = |P(W|h) - P(W|H)| = P(W|H) - P(W|h)$, the $|ASATE|$ can be represented in Figure 8A by subtracting the area corresponding to $P(W|h)$ in Figure 8B from the area corresponding to $P(W|H)$ in Figure 8A. This is what we do in Figure 10 for both horizons.

Figure 10. *2- and 4-Year Ahead Aggregate SATE (ASATE)*



Counterfactual Simulations

Medical innovations may increase health and longevity. Holding retirement age constant, such innovations make the Social Security system less sustainable. In making projections, the system must take this channel into account. Our SATE estimates can be used to provide an estimate of the countervailing effect on sustainability through increases in labor supply from higher health.

Specifically, in this section we simulate the effect of reducing in half each person's baseline likelihood of entering low health in 2 and 4 years on the population-level labor supply forecasts at those horizons. Figure 11A juxtaposes the *status quo* forecasts from Figure 7 and the simulated forecasts for the 2 years horizon. Figure 11B does the same for the 4 years case.

Halving respondents' baseline probabilities of entering low health in 2 years increases the estimated proportion of individuals predicted to work in 2 years by 2 percentage points. Halving respondents' baseline probabilities of entering low health in 4 years increases the estimated proportion of individuals predicted to work in 4 years by 3 percentage points. Note that the proportionate change in health probabilities affects more those that have significant expected health declines.

Figure 11A. *2-Year Ahead Labor Supply Forecasts, Survey vs. Simulation with* $\tilde{P}(h) = P(h)/2$

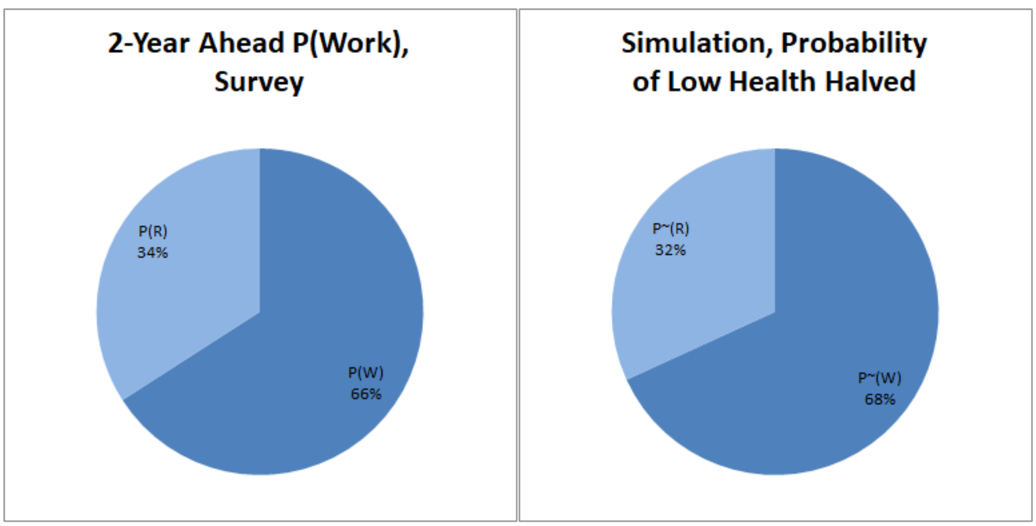
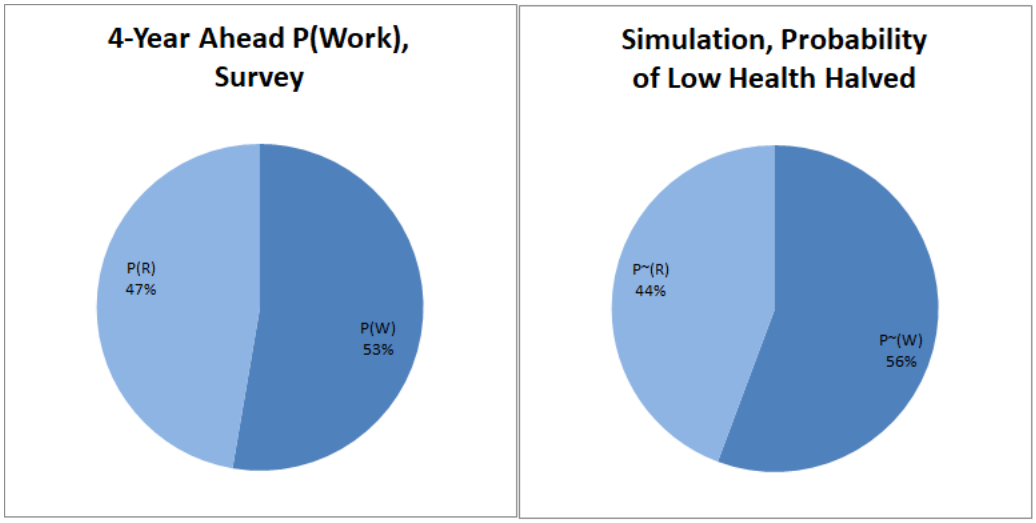


Figure 11B. *4-Year Ahead Labor Supply Forecasts, Survey vs. Simulation with* $\tilde{P}(h) = P(h)/2$



Conclusion and Next Steps

In this paper, we have provided a novel strategy for assessing the causal effect of a negative health change on the labor supply of healthy older workers, based on individuals' own estimates of their working probabilities at specified horizons under alternative scenarios about their future health. Since these effects are subjective and *ex ante* in nature, we have called them the subjective *ex ante* treatment effect (SATE). Since these effects are obtained from respondents' subjective probability of working if they enter low health minus their subjective probability of working if they remain in high health, they are individual-level effects. By aggregating these effects across individuals we have additionally derived estimates of population-level parameters, including the average SATE (ASATE) and the average subjective *ex ante* treatment effect on the treated and on the untreated (ASATT and ASATU respectively).

Using survey reports of healthy older workers participating in the Vanguard Research Initiative (VRI), we have found that respondents' labor supply expectations conditional on hypothetical health scenarios imply a zero SATE of health on labor supply at 2 and 4 years for almost 30 percent of the respondents. The remaining 70 percent reports subjective expectations which imply a strictly negative SATE (median = -40 percent and std. dev. = 24 percent at 2 years; median = -30 percent and std. dev. = 25 percent at 4 years).

Using our estimates of the 2- and 4-year ahead SATEs, we have performed two counterfactual simulations of interest to the Social Security Administration. In particular, we have simulated the effect on 2- and 4-year ahead labor supply forecasts of halving individuals' baseline probabilities of experiencing low health at those horizons. These simulations yield an increase of 2 and 3 percentage points in the estimated proportions of working individuals at 2 and 4 years respectively.

The simulation in this paper illustrates the power of the SATE approach. It allows us to simulate the effect of a change in health of the population based solely on our estimate of the SATE and the baseline health probabilities. Hence, one can analyze the likely effects of changes in health without having to rely on restrictive modeling assumptions necessary to make such inferences from estimates based on behavioral data. In particular, the SATE approach can simulate the effects of changes in policy and changes in the environment that are unprecedented in the historical data.

One can use our estimates to carry out additional analyses of interest to the Social Security Administration. First, consider policy proposals which advocate increasing the normal retirement age. In addition to the implication of such policies for payout rates of benefits, increasing the normal retirement age separately may increase working. Our estimates can be used to bound this effect based on health, by estimating bounds of labor supply at a particular age while taking into account the distribution of health. For example, to the extent that the SS early retirement age creates a norm for retirement at age 62, how much could an increase in the early retirement age to 65 potentially shift the norm?

Second, our estimates can be used as inputs into structural modeling (currently in progress). That is, instead of using realized labor supply and health histories, one can use conditional labor supply expectations as direct measures of the conditional probabilities generated by a behavioral dynamic model of labor supply.

Third, observation of health and labor supply realizations in the 2018 and 2020 administrations of the VRI survey will enable us to assess the predictive power of subjective conditional probabilities on the conditional outcomes and compare it to that of unconditional probabilities. Additionally, we plan to collect (revised) measures of 2-years-ahead work and health probabilities in the 2018 administration of the VRI.

Finally, an identical battery of questions was fielded in an experimental module of the 2016 administration of the Health and Retirement Study (HRS). Hence, we will soon be able to derive SATE estimates and use them as inputs in counterfactual policy simulations of interest to the Social Security using a more varied sample of older Americans.

References

Ameriks, John and Andrew Caplin and Minjoon Lee and Matthew D. Shapiro and Christopher Tonetti. 2015a. "The Wealth of Wealthholders." NBER Working Paper, 20972.

Ameriks, John and Joseph Briggs and Andrew Caplin and Matthew D. Shapiro and Christopher Tonetti. 2015b. "Long-Term Care Utility and Late in Life Saving." NBER Working Paper, 20973.

Arcidiacono, Peter and Joseph V. Hotz and Arnaud Maurel and Teresa Romano. 2014. "Recovering Ex Ante Returns and Preferences for Occupations using Subjective Expectations Data." NBER Working Paper, 20626.

Auerbach, Alan J. and Kerwin K. Charles and Courtney C. Coile and William Gale and Dana Goldman and Ronald Lee and Charles M. Lucas and Peter R. Orszag and Louise M. Sheiner and Bryan Tysinger and David N. Weil and Justin Wolfers and Rebecca Wong. 2017. "How the Growth Gap in Life Expectancy May Affect Retirement Benefits and Reforms." NBER Working Paper, 23329.

Banerjee, Sudipto and David Blau. 2013. "Employment Trends by Age in the United States: Why Are Older Workers Different?" MRRC Working Paper, 2013-285.

Barsky, Robert B. and F. Thomas Juster and Matthew D. Shapiro. 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *Quarterly Journal of Economics* 112(2): 537-579.

Behncke, Stefanie. 2012. "Does Retirement Trigger Ill Health?" *Health Economics* 21(3): 282-300.

Ben-Akiva, Moshe and Steven R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*, The MIT Press.

Blass, Asher A. and Saul Lach and Charles F. Manski. 2010. "Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability." *Internation Economic Review* 51(2): 421-440.

Blundell, Richard and Jack Britton and Monica Costa Dias and Eric French. 2016. The Dynamic Effects of Health on the Employment of Older Workers. MRRC Working Paper, 2016-348.

Bound, John. 1991. "Self-Reported Versus Objective Measures of Health in Retirement." *Journal of Human Resources* 26(1): 106-138.

Bound, John and Michael Schoenbaum and Todd Stinebrickner and Timothy Waidmann. 1999. "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers." *Labour Economics* 6(2): 179-202.

Bruine de Bruin, Wandi and Paul S. Fischbeck and Neil A. Stiber and Baruch Fischhoff. 2002. "What Number is "Fifty-Fifty?" Redistributing Excessive 50% Responses in Elicited Probabilities." *Risk Analysis* 22(4): 713-723.

Case, Anne and Angus Deaton. 2015. "Rising Morbidity and Mortality in Midlife Among White Non-Hispanic Americans in the 21st Century." *PNAS*, 112(49): 15078-15083.

Coe, Norma B. and Gema Zamarro. 2011. "Retirement Effects on Health in Europe." *Journal of Health Economics* 30(1): 77-86.

Coile, Courteny C. 2015. "Economic Determinants of Workers' Retirement Decisions." *Journal of Economic Surveys* 29(4): 830-853.

Delavande, Adeline and Charles F. Manski. 2015. "Using Elicited Choice Probabilities in Hypothetical Elections to Study Decisions to Vote." *Electoral Studies* 38: 28-37.

Disney, Richard and Carl Emmerson and Matthew Wakefield. 2006. "Ill Health and Retirement in Britain: A Panel Data-Based Analysis." *Journal of Health Economics* 25: 621-649.

Dwyer, Debra and Olivia Mitchell. 1999. "Health Problems as Determinants of Retirement: Are Self-Rated Measures Endogenous?" *Journal of Health Economics* 18(2): 173-193.

Fischhoff, Baruch and Wandi Bruine de Bruin. 1999. "Fifty-Fifty = 50%?" *Journal of Behavioral Decision Making* 12: 149-163.

Fisher, Gwenith G., Dorey S. Chaffee and Amanda Sonnega. 2016. "Retirement Timing: A Review and Recommendations for Future Research." *Work, Aging and Retirement* 2(2): 230-261.

Garcia-Gomez, Pilar. 2011. "Institutions, Health Shocks and Labour Market Outcomes Across Europe." *Journal of Health Economics* 30: 200-213.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80: 223-255.

Gustman, Alan L. and Olivia S. Mitchell and Thomas L. Steinmeier. 1995. "Retirement Measures in the Health and Retirement Study." *Journal of Human Resources* 30: S57-S83.

Gustman, Alan L. and Thomas L. Steinmeier and Nahid Tabatabai. 2010. *Pensions in the Health and Retirement Study*. Harvard University Press.

Hensher, David A. and John M. Rose and William H. Greene. 2005. *Applied Choice Analysis. A Primer*. Cambridge University Press.

Heckman, James J. 2005. “The Scientific Model of Causality.” Sociological Methodology 35(1): 1–97.

Hurd, Michael D. and Susan Rohwedder. 2014. “Predicting the Labor Force Participation of the Older Population.” Paper presented at SIEPR Conference on Working Longer and Retirement, October 9th and 10th, 2014, Stanford.

Jones, Andrew M. and Nigel Rice and Francesca Zantomio. 2016. “Acute Health Shocks and Labour Market Outcomes.” University of York, HEDG Working Paper, #04.

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

Kapteyn, Arie and Arthur van Soest and Julie Zissimopoulos. 2007. “Using Stated Preferences Data to Analyze Preferences for Full and Partial Retirement.” IZA Discussion Paper, 2785.

Lindeboom, Maarten and Marcel Kerkhofs. 2009. “Health and Work of the Elderly Subjective Health Measures, Reporting Errors and the Endogenous Relationship between Health and Work.” *Journal of Applied Econometrics* 24(6): 1024-1046.

Louviere, Jordan J. and David A. Hensher and Joffre D. Swait. 2000. *Stated Choice Methods. Analysis and Applications*. Cambridge University Press.

Lumsdaine, Robin and Olivia S. Mitchell. 1999. “New Developments in the Economic Analysis of Retirement.” In *Handbook of Labor Economics* volume 3, chapter 50, pp. 3261-3307.

Kapteyn, Arie and Erik Meijer. 2014. “A Comparison of Different Measures of Health and their Relation to Labor Force Transitions at Older Ages.” In *NBER Book Discoveries in the Economics of Aging*, Chicago University Press, chapter 3, pp. 115-150.

Maestas, Nicole. 2010. “Back to Work: Expectations and Realizations of Work After Retirement.” *Journal of Human Resources* 45(3): 718-747.

Manski, Charles F. 1999. “Analysis of Choice Expectations in Incomplete Scenarios.” *Journal of Risk and Uncertainty* 19(1-3): 49-66.

McGarry, Kathleen. 2004. “Health and Retirement: Do Changes in Health Affect Retirement Expectations?” *Journal of Human Resources* 39(3): 624-648.

McGeary, Kerry A. 2009. “How Do Health Shocks Influence Retirement Decisions?” *Review of Economics of the Household* 7: 307-321.

Rohwedder, Susan and Robert J. Willis. 2010. “Mental Retirement.” *Journal of Economic Perspectives* 24(1): 119-138.

Shao, Feiya. 2016. Subjective Health Shocks and Retirement Timing. Job market paper, University of Michigan.

Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.

van der Klaauw, Wilbert and Kenneth I. Wolpin. 2010. "Social Security and the Retirement and Savings Behavior of Low-Income Households." *Journal of Econometrics* 145: 21-42.

van Soest, Arthur and Hana Vonkova. 2014. "How Sensitive Are Retirement Decisions to Financial Incentives? A Stated Preference Analysis." *Journal of Applied Econometrics* 29(2): 246-264.

Wiswall, Matthew and Basit Zafar. 2016. "Human Capital Investments and Expectations about Career and Family" NBER Working Paper, 22543.

Appendix A

Table A1. *Selection of the Analytic Sample*

Selection Stages	**Sample Size**
Total sample in Survey 4	3314
Not eligible for the 2 years expectations battery	2249
Career salary reported as 0 USD	9
Not in high health	29
Inconsistent answer to 2 years expectations questions	57
2 Years Sample	**970**
Not eligible for the 4 years expectations battery	87
Inconsistent answer to 4 years expectations questions	44
4 Years Sample	**839**