Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Signature:

_____ _____

Eliska Ohler Date

Health, Health Insurance, and Retirement Expectations in the U.S.

By

Eliska Ohler
Master of Arts

Economics

David E. Frisvold, Ph.D.
Advisor

Andrew M. Francis, Ph.D.
Committee Member

Sara J. Markowitz, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Health, Health Insurance, and Retirement Expectations in the U.S.

By

Eliska Ohler
B.S., Pennsylvania State University, 2010

Advisor: David E. Frisvold, Ph.D., Vanderbilt University

An abstract of
A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Economics
2013

Abstract

Health, Health Insurance, and Retirement Expectations in the U.S.
By Eliska Ohler

This study of retirement decisions combines the previous literature and methods to estimate the joint effects of health and health insurance on individuals' retirement plans and expectations. It also examines how spouse's health and health insurance influences one's retirement plans. Using two unique variables from the Health and Retirement Study - the planned retirement age and the probability of working full-time after the age of 62 - the results were estimated both cross-sectionally and longitudinally. The results show that those who lack any early-retirement health insurance options are affected more than others by health in their retirement planning. Generally, these individuals are more likely to retire earlier only if their health worsens, while those who retain health insurance upon an early retirement tend to retire earlier regardless of the state of their health. A spouse's health does not affect retirement plans by a large amount, but some results suggest that an individual with a very sick spouse may plan to retire early – a possible evidence of care giving. Finally, this study also sheds light on why past studies reported conflicting results. It finds that the relationship between health and retirement plans may not be linear, but rather quadratic where individuals are more likely to retire early only after health worsens to a certain level.

Health, Health Insurance, and Retirement Expectations in the U.S.

By

Eliska Ohler
B.S., Pennsylvania State University, 2010

Advisor: David E. Frisvold, Ph.D., Vanderbilt University

A thesis submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Master of Arts
in Economics
2013

Acknowledgements

The author would like to thank Rachana Bhatt, Andrew Francis, David Frisvold, Sara Markowitz, Joshua Robinson, and the participants of microeconomics workshop at the Department of Economics at Emory University for invaluable input and comments.

Table of Contents

List of Tables

List of Figures

1 INTRODUCTION

In the past, many health and labor economics studies sought to answer the questions of how health and health insurance affect retirement decisions of older workers. The studies can be divided into two separate areas: (1) the effects of health on retirement and labor force transitions (Jones, Rice, and Roberts 2010; Datta Gupta and Larsen 2010; Miah and Wilcox-Gok 2007; Disney, Emmerson, and Wakefield 2006; McGarry 2004; Bound et al. 1999; Bound 1991), and (2) the effects of retiree health insurance and other types of health insurance on retirement or labor force transitions (Nyce et al. 2011; Kapur and Rogowski 2007; McGarry 2004; Blau and Gilleskie 2001; Gruber and Madrian 2002, 1995). In both cases, there are some results that contradict each other, some introduce the presence of a spouse into the model, and very few venture to examine the joint effects of health and health insurance on retirement, mostly to find insignificant results. To my knowledge, there are no studies that performed an in-depth examination of the changes in expected retirement age as affected by health, health insurance, and their interaction, even though it seems that, especially in the U.S. labor market, health status will affect retirement decisions differently, depending on the health insurance available after retirement[1].

In this paper, I propose that early-retirement health insurance availability, due to financial reasons, determines the ways in which the health of individuals and couples affects their retirement decisions[2]. I combine the two areas of labor and health economics described above and describe the way in which health affects retirement plans and how this varies across different health insurance types available to the individual and his/her spouse. I use the U.S. Health and Retirement Study's

[1]For the purpose of literary style, the terms *expected retirement age* and *planned retirement age* will be used interchangeably throughout the paper. These terms refer to the age at which an individual plans to stop working and to retire.

[2]The term *early-retirement health insurance availability* refers to the health insurance available to an individual retiring before the age 65 - the age at which one qualifies for Medicare.

(HRS) longitudinal data on retirement expectations and planned age of retirement in a fixed effects framework to test my hypothesis. In the empirical model, I address most of the current empirical problems related to estimating the effects of health on retirement, such as the occurrence of justification bias, endogeneity, reverse causality, or errors-in-variables.

With regard to data, the main innovation of this paper is the use of the planned retirement age as a dependent variable[3]. Most other studies use the actual age of retirement; however, due to many unobserved factors, the previously estimated effects of health alone on retirement may be biased. Here, I look at how the expected retirement age changes over time in a panel data framework, thus allowing for a deeper insight into retirement decisions and a better description of the forming of an actual retirement age. This may have significant policy implications in helping examine the possible effects of health insurance laws and policies, such as the Affordable Care Act, and further carefully defining the reasons for labor force detachment and attachment in later years of life.

Additional contributions of this study include: (1) looking at health and health insurance simultaneously; (2) examining a nonlinear relationship between spouse's health and retirement; (3) the use of parental data, risk aversion, and leisure proxies as control variables in order to avoid possible endogeneity; and (4) the use of all available waves of the HRS[4]. In this study, I find that for individuals who may lose health insurance upon an early retirement (before the age of 65), health matters much more in determining their retirement plans as these individuals turn to an early retirement only when their health worsens, while those who expect to retain their health insurance plan

[3]The first paper to use a panel data approach with retirement expectation was that of McGarry (2004). However, the fixed effects results were inconsistent due to the nature of the dependent variable. The use of an actual planned retirement age is advantageous in a panel data framework and gives more consistent results than the self-reported probability of working past a certain age.

[4]Kapur and Rogowski (2007) find no significance of retiree health insurance interacted with health. With regard to the relationship of health and retirement, a linear relationship has been assumed in previous studies so far. With regard to data use, most studies resort to using four waves at maximum, which greatly limits the sample size. With the exception of McGarry (2004), researchers have looked at the transitions within and out of the labor force between two periods.

to retire early regardless of their health situation. Additionally I find evidence that the relationship between health and retirement plans may not be linear, as previously assumed, but quadratic. When spouse's health is considered, I find evidence of an earlier retirement due to worsening health (possibly due to care giving). Overall, the effects of spouse's health on one's retirement are smaller in magnitude than those of one's own health.

The rest of the paper is structured as follows. Section 2 gives a brief background information on the health and health insurance characteristics of older workers. Section 3 describes the previous literature in the areas of health and health insurance and joint retirement decisions of spouses. Section 4 presents the structural models and theory of the paper. Section 5 describes the data and the variables used to construct the measures of health and the latent health stock variable. The empirical analysis can be seen in section 6, while section 7 gives a detailed review of the results of the empirical model. Section 8 concludes the study.

2 BACKGROUND

2.1 Health of Older Workers

Traditionally, those of a higher age tend to have poorer health than those who are younger. Table 1 describes the incidence of chronic and other diseases among age groups in a similar fashion to Gruber and Madrian (1996)[5]. The data shows that as an individual's age increases, the risk of having a certain medical condition increases. Although the data on health care utilization was not easily available, it is possible to infer that the utilization and health care expenditures are positively related to age. Thus, older individuals may benefit more from health insurance, which helps mitigate the

[5]Gruber and Madrian (1996) use the 1987 National Medical Expenditure Survey. Due to the age of this survey data, a new table with more recent incidence measures was created. The source of the data for table 1 is the Centers for Disease Control and Prevention.

large expenses associated with these medical conditions.

Table 1: Disease incidence by age

Age Group	20-44	25-44	45-64	44-54	55-64	65+
Self-Reported Health Status						
Good		24.00%		27.90%	29.80%	34.70%
Fair/Poor		7.20%		13.10%	19.10%	24.60%
Disease Incidence						
Stroke		0.60%		1.90%	3.80%	9.00%
Cancer		2.30%		6.60%	11.80%	23.30%
High blood pressure	10.40%		40.40%			70.10%
Emphysema		3.30%		5.90%	7.70%	10.00%
Diabetes		3.00%		9.10%	15.00%	18.80%
Heart Disease		4.80%		9.30%	16.60%	31.70%

Note: Self-reported health status, emphysema and hypertension data comes from the 2008-2010 period. The numbers represent the estimated percentage of individuals within an age group who report having a certain condition within the period of one year.

2.2 Job Lock

Because of the higher need of health insurance and the fact that the most widely used health insurance is employer-sponsored, 64% of those of age 55-64 are using such insurance plan (Jacobson, Schwartz, and Neuman 2009). This situation has created what some call a *job lock*. The term *job lock* was first used during the 1990s discussions of universal health coverage. As Jonathan Gruber defines in an article, 'Job lock refers to the fact that workers are often unwilling to leave a current job that provides health insurance for another position that might not, even if they would be more

productive in that other position' (Gruber 2009). Alternatively, in case of retirement decisions, one may choose to define a situation, in which individuals are not entering retirement because of the expensive private health insurance they would have to purchase, as a form of a job lock. Therefore, even if the marginal product of their work is less than the value that these individuals assign to their leisure time, those who expect to experience higher medical expenditures will be unwilling to retire (Gruber and Madrian 2002).

2.3 Health Insurance and Related Legislation

Up until the passage of the Patient Protection and Affordable Care Act of 2010 (PPACA), some individuals of older age struggled to find affordable health insurance. Premiums were usually very high and many plans included high deductibles. Some were denied health insurance based on pre-existing conditions. To better describe the situation, according to the U.S. Census Bureau (2012), while 67.3% and 7.1% of those of ages 45-54 were covered through an employer or by private insurance respectively in 2009, the percentage covered under these two types of plans for the 55-64 age group was 65.7% and 9.3%, with more individuals relying on Medicare or Medicaid. 13.9% remained uninsured in the 55-64 age group. This implies that there are many older workers in need of health insurance and the health insurance burden is shifting from employers to the government or private payers. To make matters worse, many were denied health insurance based on their age or pre-existing conditions. For an illustration, in 2006, the denial rate was 13% for those of age 45-49, compared to the 17%, 22% and 29% for those of age 50-54, 55-59 and 60-64 respectively (Jacobson, Schwartz, and Neuman 2009). This indicates that, for those older individuals who do not qualify for Medicaid or Medicare, group health insurance may be the only way to obtain coverage.

Attempts were made in the past at the federal level to help older and ill individuals keep or transfer their insurance upon leaving or losing a job. The Consolidated Omnibus Budget Reconcili-

ation Act (COBRA) of 1985 enabled workers who left or lost their jobs to keep the health insurance policy for themselves and any dependents for another 18 months at a 101% of its previous cost. However, this amount may be very high and possibly unaffordable for an unemployed individual. In 1996, Title I of the Health Insurance Portability and Accountability Act (HIPAA) established stricter rules on health insurance portability, favorable to workers. In the study, I control for the introduction of HIPAA although the HIPAA implementation differed among states. Theoretically, the earliest that an individual could retire, while keeping his or her current employer's insurance until reaching the Medicare eligibility age, would be at the age of 63.5.

When individuals in the U.S. are of poor health, which does not allow them to work and the condition is terminal or long-term, they may qualify for Medicare two years after the date of entitlement to the Social Security Disability Insurance (SSDI). Finally, some individuals of income below a certain Federal Poverty Level threshold may qualify for Medicaid.

3 LITERATURE REVIEW

3.1 Health and Retirement

Over the past few decades, researchers have been trying to determine the effects of health on retirement decisions of individuals. This question has been addressed from many different angles using many different techniques and measures. Some studies find that those who report poor health tend to retire earlier (Jones, Rice, and Roberts 2010; Datta Gupta and Larsen 2010; Disney, Emmerson, and Wakefield 2006; McGarry 2004; Bound et al. 1999). However, others find that this is true for only later onsets of illnesses and that those who had experienced poor health prior to old age tend to be either less affected by their health (Bound et al. 1999) or even tend to retire later (Miah and Wilcox-Gok 2007). Some find different effects for different illnesses (Datta Gupta and Larsen

2010). Green (2006) shows in her study that among those who continue working after the age of 65, those of poor health tend to work more than those of better health. This is also implied by Bound et al. (1999), who find that contemporaneous decline in health may result in an earlier retirement, while persistent poor health may result in a later retirement.

The findings above suggest that this topic should be examined more thoroughly and that one needs to be careful as to avoid potential endogeneity and measurement error biases (Datta Gupta and Larsen 2010). Furthermore, as mentioned later in section 3.2, the type of health insurance that one has influences retirement decisions, especially in the U.S.. The studies in this area frequently use European data from countries such as Britain (Disney, Emmerson, and Wakefield 2006; Jones, Rice, and Roberts 2010), Denmark (Datta Gupta and Larsen 2010), and Netherlands (Lindeboom and Kerkhofs 2009). Since European nations tend to be much more generous in their disability, early-retirement, and retirement benefits and most have universal health insurance coverage, their results may not be representative of the U.S., where employer-sponsored health insurance plans are the main source of health insurance before the age of 65 and where many older workers are uninsured.

Among those who study retirement behavior in the U.S., the Health and Retirement Study (HRS) is the most prevalent source of data. Bound et al. (1999) use Waves 1, 2, and 3 of the HRS to analyze, separately by gender, the effects of both contemporaneous and persistent health conditions on labor force transitions. They use a constructed latent variable as an instrument and examine a multitude of labor force transitions such as exit via SSDI or finding a different job. However, the authors do not use panel data, but rather concentrate on changes from Wave 2 to Wave 3, using Wave 1 as a source of lagged values. With respect to health, they conclude that those of poorer contemporaneous health are more likely to exit the labor force, while poor persistent (lagged) health is associated with continued labor force attachment. Disney, Emmerson, and Wakefield (2006)

conducted a similar study using British panel data and found, unlike Bound et al., that both lagged and current health measures affect labor force exits in the same direction - individuals of poorer health are more likely to exit.

In 2004, McGarry published a study of the impact of health on retirement expectations. The author used the first two waves of HRS data to examine the variation in self-reported probabilities of working full-time after the age of 62. By doing this, the author claims to avoid the justification bias, which complicated other studies that used a binary indicator variable of retirement[6]. The results of this study, found using cross-sectional or simple differences model, suggest that poor health leads to a decreased probability of working after the age of 62.

In terms of econometric analysis, the most detailed analysis method to date was described by Jones, Rice, and Roberts (2010), who looked at the effects of self-reported health on early retirement, using 12 years of British data and doing a separate analysis by gender of the effects of health on retirement. They combine many of the previous methods to come up with a complex econometric analysis due to the many possible problems (later described in section 3.4). The authors use a latent variable model, in addition to lagged health measures and hazard models, and find that ailing health increases the probability of early retirement.

3.2 Health Insurance and Retirement

Many scholars in the past had examined the relationship between health insurance and retirement, both in reduced and structural forms. They generally confirm that the threat of losing health insurance may deter individuals from retiring (Gruber and Madrian 2002). In many reduced form studies, the availability of retiree health insurance has been a popular explanatory indicator variable. In general, researchers find that the availability of retiree health insurance does increase the probability of

[6]Some of the most recent studies include Jones, Rice, and Roberts (2010), Disney, Emmerson, and Wakefield (2006) and Bound et al. (1999).

early retirement (Nyce et al. 2011; McGarry 2004; Blau and Gilleskie 2001; Gruber and Madrian 2002, 1995). By contrast, when Kapur and Rogowski (2007) analyzed this issue in the context of joint retirement of couples, they found only a modest effect of retiree health insurance on retirement patterns, with the effect being stronger for women as compared to men. Additionally, the authors found no significance of interaction terms of health with health insurance.

3.3 Partner's Health and Retirement

With regard to partner's health affecting one's retirement decisions, only a few pieces of evidence can be gathered from previous research. For example, a theoretical discussion of joint retirement decisions was done by Blau and Gilleskie (2006), who model and later empirically examine these decisions. Another study, by Madrian and Beaulieu (1998), concludes that the Medicare eligibility of a spouse does affect one's retirement age, which may suggest that a partner's health and costs of care are considered in retirement decision making. Not completely unrelated to health, Jones, Rice, and Roberts (2010) find that having a working partner decreases the hazard of early retirement, but that there are no effects of a spouse's or a partner's health on early retirement. However, the authors used British data where health insurance dependence or the lack of health insurance is not present as widely as in the U.S. Finally, using data from the HRS, Kapur and Rogowski (2007) find that spouse's health problems may result in an early retirement, possibly due to care-giving.

3.4 Potential Problems and Biases

As all authors mention in their work, the topic of health and retirement is confronted by many empirical and data issues (Lindeboom and Kerkhofs 2009; Datta Gupta and Larsen 2010; Jones, Rice, and Roberts 2010; Bound 1991). Jones, Rice, and Roberts (2010, p. 867), divide the possible sources of endogeneity into two groups: Type I, due to unobservables and possible simultaneity;

and Type II, due to justification bias and classical errors. Lindeboom and Kerkhofs (2009) address *simultaneity* and *justification bias* in their recent work. In case of *simultaneity*, they mention that those who worked longer in the past may have worse health than those who worked less, while health may affect how much and how long one works. Additionally, *justification bias* may appear in studies using only self-reported subjective health data because individuals may justify their intentions to retire early by poor health (an acceptable excuse); thus making their self-reported health status dependent on their employment status. The authors address both issues in their work by using a three stage model and including lagged health measures. Jones, Rice, and Roberts (2010) find that once lagged health measures or health shocks are included, the justification bias becomes a smaller issue and at least one study has found that the justification bias may not be present in the U.S. self-reported health data (Dwyer and Mitchell 1999).

Bound (1991) and subsequently (Disney, Emmerson, and Wakefield 2006) address the *endogeneity* and *errors-in-variables* bias by constructing a latent variable - health stock - which is a predicted health status based on current health conditions and other characteristics of the respondent. Bound (1991) points out that objective health measures may suffer from a classical measurement error (which may bias the coefficients towards zero) and argues therefore for using a subjective health status, claiming that endogeneity (or justification bias) and the errors-in-variables bias may work in opposite directions. In a later paper, Bound et al. state that using self-reported more specific objective measures of health in the HRS may lead to a lesser endogeneity. Datta Gupta and Larsen (2010) counter Bound's argument saying that both subjective and objective self-reported health measures are erroneous and use matched medical records to avoid bias. Finally, Datta Gupta and Larsen find that for some illnesses, self-reported measures do not matter, while they do matter for others[7].

[7]See Datta Gupta and Larsen (2010, p. 809) for detailed comparisons of predicted retirement plans using both official medical records and self-reported data.

When it comes to health insurance choices, another host of potential issues may appear. French and Jones (2011) find some unobserved differences between workers with different health insurance plan types. Those with more generous plans and with retiree health insurance may have stronger preferences for leisure and thus may also choose to retire early in addition to choosing the retiree health insurance; however, some point out that few individuals choose their job solemnly on the basis of retiree health insurance availability(Madrian, Burtless, and Gruber 1994). I would add that risk aversion may also affect both health insurance choices and retirement plans; in some cases, it may also affect health. Later, in section 6.3, these biases are addressed in the context of this study.

4 CONCEPTUAL FRAMEWORK

As many economists have shown before, individuals and households maximize their expected utility subject to their budget constraint. The result of such utility maximization is in many cases the determination of their retirement date R (Disney, Emmerson, and Wakefield 2006; French and Jones 2011). Equation 1 displays this relationship in an intertemporal labor supply model (Bound et al. 1999; Blau and Gilleskie 2006; French and Jones 2011). The expected utility U may be affected by many factors such as joint consumption C_j, leisure L_j, and other factors Z_j which include health h_j but also other personal characteristics and tastes X_j.

$$max\, E_t \sum_{j=t}^{T} \beta^{j-t} * U(C_j, L_j, Z_j\,[h_j, X_j]) \tag{1}$$

The expected utility function is subject to a budget constraint, equation 2, where W_j is wage (dependent on health) multiplied by hours of work H_j[8]. C_j represents consumption and M_j represents medical expenses, which we can assume are dependent on health and a proportional out-

[8]Bound et al. (1999) suggest that health may determine the type of job that one may perform and thus the wage.

of-pocket expenditure. A_j represents an individual's assets; and r_j the interest rate (all at period j).

$$A_{j+1} = (W_j * H_j - C_j - M_j) + (1 + r_{j+1}) * A_j \tag{2}$$

However, after a certain time period k, once an individuals or a couple retires, W becomes zero, and savings A (including Social Security income) become the only source of funds, while consumption is dependent on similar variables as before (see equation 3). The value of k stands for the time which elapses between period j and R, the actual retirement date.

$$A_{j+k+1} = (1 + r_{j+k+1}) * A_{j+k} - C_{j+k} - M_{j+k} \tag{3}$$

Health may affect R, or k, in multiple ways. On one hand, higher expected health expenditures M may induce individuals to work longer. Risk averse individuals may be strongly attached to their employers – a source of group insurance – in order to smooth consumption and maximize their expected utility. On the other hand, some illnesses or health conditions make work more costly or impossible. Additionally, lower life expectancy T due to an illness may lead to an earlier retirement because available accumulated wealth A is now divided into fewer years and thus relatively rises. Changes in a partner's health may lead to similar effects, as individuals may be planning to retire jointly[9]. Additionally, poor health of a spouse may lead to earlier retirement in order to provide care for one's spouse.

However, the way in which R is affected by health may depend on the accumulated assets and on expected health care consumption. Health insurance in the U.S. is not provided uniformly

[9]For example, Hurd (1990) states that as much as couples plan jointly in the early stages of their life, they also plan jointly their retirement. Factors that may affect the degree of retirement coordination include the enjoyment of time spent together, family assets, Social Security eligibility, and the age difference of partners.

by the government but rather by employers, through a spouse's employer or through an individual policy with many individuals left uninsured[10]. The type of health insurance affects the expected cost of health care consumption. If one is uninsured, health care consumption may lead to a larger asset deterioration and higher expected costs than if one is insured. Additionally, if one is not eligible for Medicaid (due to income above the threshold level) or Medicare (either due to low age or the lack of a significant disability), it will be in one's interest to keep any current health insurance policy in existence in order to maximize long term consumption and life expectancy since having health insurance in the case of low savings may be the only way to ensure access to quality care. This may then lead individuals to a later retirement.

Similar logic can be applied to a case, where a spouse/partner, whose health recently deteriorated, were dependent on an individual's health insurance. It may be, for example, in the husband's interest to provide health insurance for his wife until she turns 65 and qualifies for Medicare, because the husband maximizes his expected utility by maximizing the consumption of time spent with his wife (this, of course, depends on whether he gains utility from doing so)[11]. Especially in times of recession, when the accumulated assets of many older workers lost much of their value, the availability of health insurance may affect how health impacts the retirement date R. This theory contradicts the findings of Kapur and Rogowski (2007), perhaps because the relationship between spouse's health and retirement may be nonlinear; that is, until a spouse's health deteriorates to the point of needing care, one may choose to retain a job in order to provide the spouse with health

[10]To qualify for Medicare before the age of 65, one must qualify for the Social Security Disability Insurance (SSDI). The criteria is very strict and means tested, and only those who are found with severe conditions (unable to walk, sit, or remember for example) will qualify. Additionally, even if one qualifies for SSDI, it may take up to two years to qualify for Medicare. To qualify for Medicaid, one needs to pass the means test with income below a certain threshold. Because of COBRA, workers who leave their jobs are also able to keep their old insurance policy at 102% of the old premium, but these policies may be very costly, especially to the jobless (Scandlen 2001).

[11]Madrian and Beaulieu (1998) find that spouses' Medicare eligibility affects retirement decisions.

insurance.

5 DATA

In order to test the theory outlined in section 4, I will be using the biennial Health and Retirement Study (HRS), prepared by the RAND Center for the Study of Aging. Unlike previous studies, I intend to make a full use of all the available data in the HRS, ranging from 1992 to 2010. The HRS is suitable for this kind of study because it collects data on health, income, retirement, family, leisure, work, and other characteristics for a large sample of individuals of age 50 and above and their spouses[12]. This is a longitudinal study and individuals are asked a similar set of questions every year. In addition to using the HRS, some data for this study were obtained using the Life Tables and Deaths – final and preliminary – data published by the National Center for Health Statistics, local area unemployment statistics from the Bureau of Labor Statistics, seasonally adjusted annual consumer price index (CPI) provided by the Federal Reserve Bank of St. Louis, and full retirement age by year of birth data published by the Social Security Administration.

5.1 Data Processing

In the process of cleaning the data, the HRS Version L dataset was converted into a longitudinal dataset. Next, a few specific variables from HRS Fat Files were merged with this original dataset. Data for annual unemployment rates in Census divisions, expected age of death, social security eligibility age, and unemployment were added for both the respondent and the respondent's spouse[13]. Some data that should not vary across waves, but were found missing in some waves (such as gender

[12] Age 50 as of the first recording time of each cohort. See Health and Retirement Study (2011) for a complete listing of the number of respondents and response rates by HRS waves and cohorts.

[13] Unfortunately, the publicly available HRS data do not include the state of residence and thus do not allow to control for state-specific characteristics.

or the date of birth), were filled in using a statistical software. Employment and marital status were imputed if the individual had not changed an employer or a spouse respectively. All dollar amounts (such as medical expenditures, annual income, assets, etc.) were discounted using the CPI to their 1992 values in order to allow for accurate comparisons among individuals and across time.

In order to answer the research question at hand using the fixed effects framework, the HRS dataset was narrowed down to include only individuals from the Original HRS, War Babies (WB), and Early Baby Boomers (EBB) cohorts. The whole sample included 13,525 Original HRS, 2,760 WB, and 3,522 EBB individuals - a total of 19,807 respondents or their spouses. Next, I eliminated 800 individuals eligible for or actively receiving Medicare or Social Security Benefits (whose age at the time of the HRS interview was at or above 65 years)[14]. Additionally, the sample was restricted to those who were working part-time, full-time or were unemployed (looking for a full-time or part-time job) at the time of the interview, excluding 4,560 respondents[15]. I also restrict the sample to those who were not receiving either Social Security Disability Insurance (SSDI) or Supplemental Security Income (SSI), similarly to Bound et al. (1999), eliminating 75 respondents. The final sample consists of 14,372 individuals, who constitute approximately 75.61% of those with same cohort and age characteristics. Finally, for panel data analysis, I restrict the sample to those for whom I have observations for more than one wave, which reduces the sample by 2,562 individuals. When looking at the retirement expectations of married couples, I additionally omit 2,200 individuals who are not married.

[14]This is because those who are eligible for Medicare or Social Security Benefits may have very different retirement expectations, and thus may react very differently to changes in health or health insurance.

[15]Unlike McGarry (2004), I do include those who are self-employed.

5.2 Dependent Variables

The two variables in the HRS that allow for a longitudinal analysis of retirement expectations, used in this study are 1) the planned retirement age *R*; and 2) the self-reported probability of working past the age of 62 *P(62)*[16]. Two variables, instead of one, are used to ensure robust results. Figures 1 and 2 show the histograms of the two variables, which is helpful for the purposes of empirical analysis. Table 2 displays the means of the dependent variables for survey years (waves) for different cohorts. Questions about the variation of the dependent variables for individuals over time are addressed in appendix A, which shows the distribution of the differences between the highest and the lowest values of the dependent variables by respondent and also gives some reasons for the differences between the planned and actual retirement age. Tables 3 and 4 display the mean planned retirement age and the probabilities of working full-time after the age of 62 by self-reported health status of both the respondent and his or her spouse.

---

[16]For the analysis of planned age of retirement, the dependent variable is available for 49.4% observations. The planned age of retirement is missing when an individual does not plan to stop working or when the interview is conducted via a proxy.

The questions in the survey specifically asked (in order of mention):

1. *At what age do you plan to stop working?*
2. *Thinking about work generally and not just your present job, what do you think are the chances that you will be working full-time after you reach age 62?*

Although I intended to use the probability of working full-time after the age of 65 as a dependent variable, because of the fact that the majority of individuals reported this probability to be zero and that those who did not would be a very specific group, I decided not to include this variable in the analysis.

The benefits of using continuous rather than binary measures of labor force attachment are described in McGarry (2004).

With regard to the choice of these two dependent variables, the actual age of retirement was not chosen as it did not allow for a longitudinal analysis. The deviation from retirement plans was not studied in this paper, but should be a subject to further research. A preliminary summary of its determinants can be found in table 3 of appendix A.

Table 2: Means of the dependent variables by year and cohort

	Planned retirement age			Probability of F-T work after 62		
Cohort	Original HRS	WB	EBB	Original HRS	WB	EBB
1992	62.927			0.488		
1994	62.553			0.454		
1996	63.096			0.469		
1998	63.317	62.682		0.476	0.471	
2000	63.842	62.783		0.535	0.484	
2002	64.583	63.577		0.557	0.496	
2004	65.115	64.296	63.731	-	0.540	0.513
2006	65.491	64.984	64.126	-	0.579	0.532
2008	-	65.411	64.400	-	0.631	0.601
2010	-	66.447	65.556	-	0.621	0.585

Table 3: Means of planned retirement age by self-reported health status

Respondent's health / Spouse's health	Excellent	Very good	Good	Fair	Poor
Excellent	63.362	62.954	62.900	61.965	60.288
Very good	63.033	63.014	62.671	62.137	60.360
Good	63.040	62.987	62.441	61.830	61.642
Fair	62.032	62.017	61.949	61.230	62.417
Poor	61.570	61.202	61.255	61.192	60.995

Table 4: Means of *P(62)* by self-reported health status

Respondent's health / Spouse's health	Excellent	Very good	Good	Fair	Poor
Excellent	0.541	0.500	0.478	0.412	0.369
Very good	0.516	0.492	0.475	0.426	0.423
Good	0.538	0.498	0.468	0.439	0.440
Fair	0.555	0.504	0.475	0.476	0.472
Poor	0.543	0.585	0.519	0.445	0.483

Figure 1: Histogram of planned retirement age

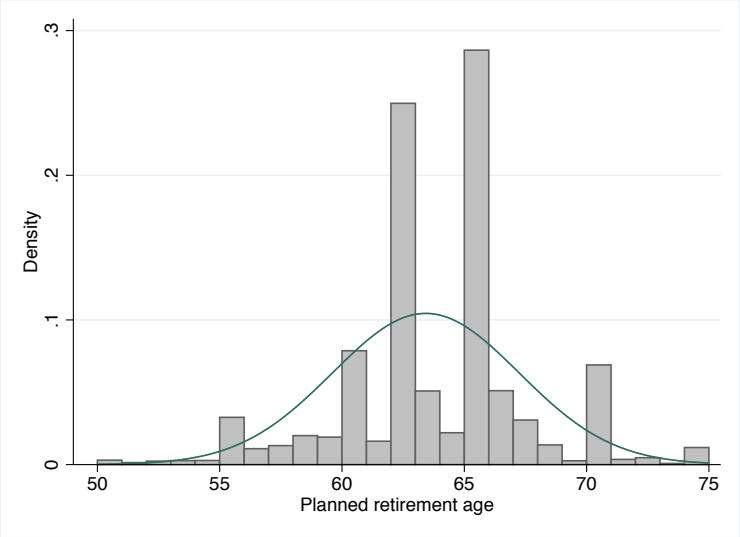
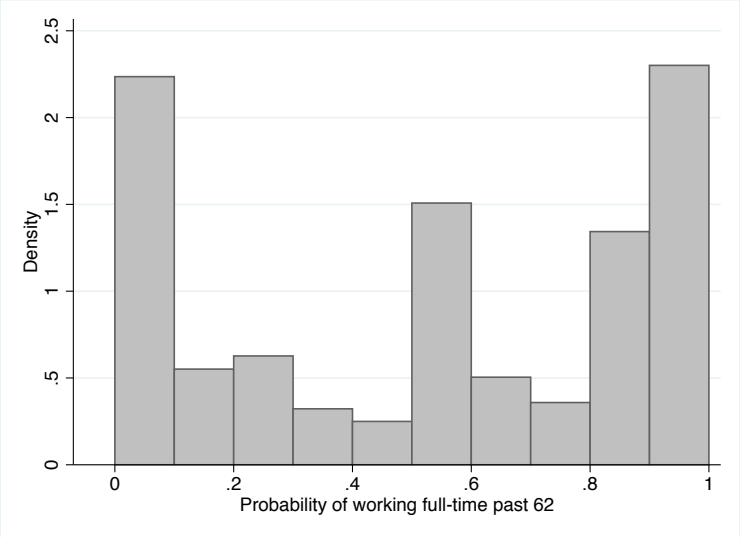


Figure 2: Histogram of the probability of working full-time after the age of 62



5.3 Health Variables

Although Bound et al. (1999) make a case for using lagged health measures to see the effects of persistent or past health conditions, I will focus mainly on contemporaneous measures in the panel data analysis context. This is because I am interested in the marginal effects of health on retirement plans. The main explanatory health variables used in this study could be classified into four categories – self-reported health, medical care use, illness occurrence, and medical expenditures.

However, self-reported health status will be the focus point of the study, while the other variables will be used to construct the latent health status[17]. Table 5 displays the means for each health variable for the respondents overall and also by gender. Females in the sample tend to have higher difficulties with activities of daily living, but they tend to spend less time in hospitals and nursing homes. Males seems to visit doctors and dentists less and have lower prevalence of cancer, lung disease, psychological problems and arthritis.

[17]The question regarding self-reported health status and asked in the interview was, *'Would you say your health is excellent, very good, good, fair, or poor?'*.

Table 5: Means of the health variables of interest

	All	Male	Female
Self-reported health			
Self-reported health status	2.401	2.411	2.393
	(1.006)	(1.016)	(0.998)
Activities of daily living score	0.061	0.057	0.064
	(0.330)	(0.317)	(0.340)
Walking across a room	0.009	0.008	0.011
	(0.096)	(0.087)	(0.103)
Walking a block	0.033	0.027	0.038
	(0.178)	(0.161)	(0.190)
Walking several blocks	0.118	0.095	0.136
	(0.322)	(0.293)	(0.343)
Getting in and out of bed	0.023	0.020	0.025
	(0.149)	(0.139)	(0.156)
Sitting for two hours	0.152	0.136	0.166
	(0.359)	(0.343)	(0.372)
Getting up from a chair	0.257	0.215	0.291
	(0.437)	(0.411)	(0.454)
Walking up a flight of stairs	0.066	0.045	0.082
	(0.248)	(0.208)	(0.275)
Walking up several flights of stairs	0.311	0.228	0.380
	(0.463)	(0.419)	(0.485)
Lifting 10 pounds	0.103	0.049	0.148
	(0.304)	(0.216)	(0.355)
Picking up a small object (dime)	0.024	0.022	0.027
	(0.154)	(0.147)	(0.161)
Negative feelings score	1.175	1.004	1.297
	(1.761)	(1.590)	(1.865)
Probability of living past age of 75	0.672	0.651	0.689
	(0.272)	(0.280)	(0.266)
Health limits work	0.092	0.090	0.093
	(0.289)	(0.287)	(0.290)
Medical care use			
Nights in hospital	0.682	0.736	0.638
	(4.445)	(3.807)	(4.908)
Nights in nursing home	0.066	0.138	0.007
	(7.916)	(11.770)	(0.362)
Doctor visits	6.300	5.313	7.115
	(11.192)	(9.760)	(12.188)
Prescription drugs	0.607	0.527	0.670
	(0.489)	(0.499)	(0.470)
Outpatient surgery	0.166	0.159	0.172
	(0.372)	(0.365)	(0.378)
Dentist services	0.721	0.688	0.748
	(0.448)	(0.464)	(0.434)
Special facilities	0.046	0.042	0.049
	(0.209)	(0.201)	(0.215)

Continued on next page

Continued from previous page

	All	Male	Female
Illness occurrence			
Has high blood pressure	0.375	0.395	0.359
	(0.484)	(0.489)	(0.480)
Has diabetes	0.102	0.115	0.091
	(0.302)	(0.319)	(0.287)
Has cancer	0.058	0.040	0.072
	(0.233)	(0.195)	(0.259)
Has lung disease	0.049	0.043	0.053
	(0.215)	(0.203)	(0.224)
Has heart problems	0.088	0.106	0.073
	(0.283)	(0.308)	(0.259)
Has had a stroke	0.016	0.019	0.014
	(0.126)	(0.135)	(0.119)
Has psychological problems	0.104	0.070	0.132
	(0.305)	(0.256)	(0.338)
Has arthritis	0.373	0.319	0.417
	(0.484)	(0.466)	(0.493)
Medical expenditures			
Total medical OOP exp. (in $1,000)	1.897	1.671	2.083
	(5.070)	(5.028)	(5.097)
Total medical exp. (in $1,000)	6.011	6.128	5.912
	(20.497)	(22.081)	(19.041)

Mean coefficients with standard deviations in parentheses. Number of observations may differ depending on the variable.

Appendix B gives the definitions of the health variables mentioned above and shows the means of health measures by self-reported health status. The means show that the self-reported health status reflects all other health variables and that those with worse health status tend to have more illnesses, higher medical expenditures, use more medical care, and have higher difficulties performing daily tasks.

5.4 Health Insurance

When it comes to current health insurance of respondents — that is, the one available at the time of interviews, not after an early retirement — some individuals have more than one health insurance source. There are also several ways of looking at one's health insurance. First, we can divide individuals according to the health insurance type at the time of the interview. We may also group individuals by whether or not their employer offers a retiree health insurance option — a health insurance at a lower or discounted rate until retirement. This health insurance differentiation has been very popular among economists (Nyce et al. 2011; McGarry 2004; Blau and Gilleskie 2001; Gruber and Madrian 2002, 1995; Kapur and Rogowski 2007). Finally, we may distinguish individuals according to the availability of health insurance in case of retirement before the age of 65 (see appendix C for explanation and reasoning).

In this study, in order to test for the robustness of results, I will be using the second and third method of distinguishing among health insurance options. Appendix C defines these health insurance types and gives the means of the main health insurance variables by self-reported health status. Table 6 displays the summary statistics for the selected health insurance variables[18]. Overall,

[18]For those where both the respondent and his/her spouse had health insurance through an employer, but no retiree insurance available, I made the assumption that the respondent, in case of an early retirement, could obtain health insurance through spouse. Additionally, approximately 31% of sample respondents reported having more than one source of health insurance. The vast majority was due to the fact that individuals were insured through their own employer in addition to being insured through a spouse. This fact does not seem to alter the estimation in any way since all these

it seems that females tend to be disadvantaged when it comes to health insurance.

Table 6: Means of the health insurance variables of interest

	All	Male	Female
Division by the availability of retiree health insurance			
Retiree	0.304	0.357	0.261
	(0.460)	(0.479)	(0.439)
Employer but no retiree	0.311	0.318	0.306
	(0.463)	(0.466)	(0.461)
Government-provided	0.049	0.056	0.043
	(0.216)	(0.231)	(0.202)
Other	0.355	0.296	0.404
	(0.479)	(0.456)	(0.491)
Division by the health insurance type in case of early retirement (no spouse)			
Single — insured	0.567	0.660	0.491
	(0.495)	(0.474)	(0.500)
Single — uninsured	0.433	0.340	0.509
	(0.495)	(0.474)	(0.500)
Division by the health insurance type in case of early retirement (married)			
Married — insured	0.484	0.596	0.391
	(0.500)	(0.491)	(0.488)
Married — insured/uninsured	0.084	0.063	0.100
	(0.277)	(0.244)	(0.300)
Married — uninsured/insured	0.195	0.133	0.246
	(0.396)	(0.340)	(0.430)
Married — uninsured	0.238	0.207	0.263
	(0.426)	(0.405)	(0.440)
Observations	50933	50933	

Mean coefficients with standard deviations in parentheses.

Table 9 in appendix C displays the means of health measures by the current (at the time of the interview) health insurance type of the respondent. The descriptive statistics show that those with retiree health insurance available seem to be of similar health (when judging by illness occurrence) to others, but tend to report better health status and abilities compared to other groups, while having similar health care use and expenditures. Otherwise, there is no clear pattern among the other three groups. Table 10 in appendix C displays the means of health measures by post-early-retirement health insurance options. In this table, it seems that those who would remain uninsured (columns 3

individuals would be counted as insured now by their employer and are properly accounted for in the health insurance groups in case of an early retirement. Finally, roughly 43% of respondents change health insurance over the time of their being in the HRS.

and 4) after an early retirement have the worst health of the four groups and have the highest current use of medical care.

5.5 Control Variables

Appendix D lists the control variables used in this analysis. The means of the main control variables are displayed in table 7[19]. Because I am combining observations from different years and cohorts, I include controls for cohort, year, real GDP, and regional unemployment rate[20]. All dollar values were discounted to their 1992 value, as mentioned in section 5.1. Furthermore, appendix D displays the means of these control variables compared across health insurance types (both current and early-retirement) in order to show the differences in personal and other characteristics among those with different insurance types. The tables show that those with worse health insurance situation (present or future) are more likely to be female, Hispanic, less educated, work part time or be unemployed, have lower income, work for a smaller employer with a lower probability of being covered with a union contract, and tend to be without an employer-provided pension. It seems that overall, those with better insurance fit more the description of white collar workers as they tend to have the highest income and weekly hours of work and tend to have more stressful, but less physically demanding jobs.

[19]The variables were chosen to be comparable to table 1 in McGarry (2004, p. 631). Additionally, I do not include pension wealth and household wealth separately, but rather use net household wealth. This measure includes all savings and assets less all debts and mortgages. It is very important that debts are taken into consideration when looking at assets and savings and neither pension, nor household wealth enable us to do so.

[20]The summary statistics for these measures are not displayed, but are available upon request.

Table 7: Means and standard deviations of the main control variables

	Mean	Std. Deviation
Demographic		
Age at interview	56.174	5.133
Gender (1 - female, 0 - male)	0.548	0.498
Individual is white/Caucasian	0.800	0.400
Individual is black/African American	0.145	0.352
Individial is Hispanic	0.028	0.165
Individual is of other race/ethnicity	0.027	0.163
Years of education	12.969	2.945
Individual is protestant	0.618	0.486
Individual is catholic	0.272	0.445
Individual is Jewish	0.020	0.140
Probability of living past age of 75	0.672	0.272
Marital status	0.791	0.407
Total number of children	3.114	1.993
Income risk aversion	3.265	1.368
Individual has good friends in the neighborhood	0.606	0.489
Parental information		
Mother is alive	0.445	0.497
Father is alive	0.208	0.406
Employment		
Works full-time	0.811	0.392
Works part-time	0.156	0.363
Unemployed	0.033	0.178
Veteran status	0.207	0.405
Log of annual income	3.266	0.981
Average weekly hours of work	41.329	13.533
Experience	32.518	10.330
Employer size	337.864	1521.729
Occupation	5.967	4.771
Industry	7.941	3.831
Union contract	0.192	0.394
Job requires a lot of physical effort	2.248	1.126
Job involves a lot of stress	2.802	0.816
Individual enjoys work	2.928	1.201
Individual finds salary adequate	2.781	0.775
Reported staying in current job because of health insurance	0.323	0.468
Weeks of paid vacation	3.311	4.670

Continued on next page

Continued from previous page

	Mean	Std. Deviation
Marriage		
Couple household	0.791	0.406
Length of current marriage	27.016	11.547
Time spent with spouse is enjoyable	3.118	0.716
Free time is spent together with spouse	2.398	0.734
Retirement funds		
Net household wealth (in $1,000)	341.737	1247.141
Has pension	0.586	0.492
Previously offered retirement incentive	0.036	0.186
Observations	50933	

6 MODEL

In order to properly estimate and identify the effects of health and health insurance on retirement, there are multiple steps needed to be done in order to avoid the potential biases caused by measurement errors and endogeneity. To summarize the process, first, a latent variable will be created to be used as an instrumented variable for the error-filled health status (Bound et al. 1999, p. 186); second, a fixed effects or pooled model will be used to estimate the coefficients on the correct self-reported health status (latent health status), health insurance, and their interaction.

6.1 First Step: Latent Variable Estimation

The self-reported health status is one of the main variables of interest, yet also the one most prone to misreporting, which may lead to a bias of estimates (Datta Gupta and Larsen 2010) and the magnification of potential identification problems caused by errors-in-variables in the fixed effects framework (Griliches and Hausman 1986). For these reasons, I follow what was suggested in previous literature and create the latent counterpart of the self-reported health status[21]. Additionally, I also compute the probability of poor or fair health as a latent variable.

Similarly to what has been already described by Bound et al. (1999), the "true health", η_{it} of an individual i at time t can be described with equation 4, where x_{it} represents exogenous personal characteristics such as age, education, region, etc., z_{it} is a vector of health characteristics such as current illnesses, and ν_{it} represents unobserved characteristics – aspects of health not captured by x_{it} and z_{it}.

[21] Bound et al. (1999) and Disney, Emmerson, and Wakefield (2006) suggest that the self-reported health status should be replaced with its error-ridden proxy or the "health stock" - latent health status.

$$\eta_{it} = x'_{it}\delta_t + z'_{it}\gamma_t + \nu_{it} \tag{4}$$

However, individuals do not report this true health, but rather a subjective health status, h_{it}, on a scale from 1 to 5 (see appendix B). This subjective health report's latent counterpart then becomes h^*_{it}, which is a function of η_{it} and a reporting error ε_{it}, which is assumed to be uncorrelated with ν_{it} [22]. Equation 5 depicts this relationship.

$$h^*_{it} = \eta_{it} + \varepsilon_{it} \tag{5}$$

Thus, combining equations 4 and 5, we get the following:

$$\begin{aligned} h^*_{it} &= x'_{it}\delta_t + z'_{it}\gamma_t + \nu_{it} + \varepsilon_{it} \\ &= x'_{it}\delta_t + z'_{it}\gamma_t + u_{it}, \end{aligned} \tag{6}$$

which can be estimated in an ordered probit regression, under the assumption of normality of u_{it}. $\hat{h}^*_{it}$, the fitted value of h^*_{it}, will be used as an error-free replacement of the subjective self-reported health status in the model. As Disney, Emmerson, and Wakefield (2006) and other authors have noted before, an important assumption is that ε_{it} is not correlated with any variables used in estimating the fixed effects model described in equations 8 and 9, which the authors find to be mostly true. Similarly to Disney, Emmerson, and Wakefield, I estimate and fit these equations for each year of observations separately. Appendix E explains in detail how the fitted values were obtained.

[22] Overall, most literature assumes this to be true as it is difficult to think of unobserved characteristics that would affect both one's health and one's reporting error. It would be, for example, difficult to conclude whether those of better or worse health would have a lower or higher reporting error.

6.2 Second Step: Pooled Model and Fixed Effects Estimation

Having estimated the latent health status, I first use a simple pooled linear model (equation 7) and later expand the analysis to include the fixed-effects model (equations 8 and 9) to allow for time-invariant endogeneity, which can be caused by unobserved characteristics of individuals or by leisure preferences and risk aversion that could have arisen from former experiences or cultural upbringing.

As mentioned before, the first estimation step was employed to avoid the errors-in-variables bias in case of self-reported health measures. In case of health insurance measures, the problem of endogeneity (and also partially self-selection) due to unobserved leisure preferences or risk aversion has been mentioned in many past studies such as the one by Nyce et al. (2011). I later expand the analysis to control for the time variant portion of these two variables by controlling for income risk aversion and a few leisure characteristics. Additionally, I include interactions of self-reported health status and the health insurance type of an individual. Finally, I include a squared health status in the model to test the hypothesis that the relationship of health status and labor force exit is non-linear. With regard to control variables, the fixed effects model identifies only the marginal effects of time-varying variables, thus the effects of time-invariant variables are unknown. This will not create problems in the estimation, because all of the variables of interest vary with time[23].

$$Exp_i = \beta_0 + \beta_1 * \hat{h}_i + \beta_2 * hi_i + \beta_3 * \hat{h}_i * hi_i + \sum_m \gamma_m * x_{m,i} + \varepsilon_i \tag{7}$$

$$Exp_{it} = \beta_0 + \beta_1 * \hat{h}_{it} + \beta_2 * hi_{it} + \beta_3 * \hat{h}_{it} * hi_{it} + \sum_m \gamma_m * x_{m,it} + \phi_t + \tau_i + \varepsilon \tag{8}$$

[23] When time-invariant control variables were interacted with time indicators, the results remained unchanged. Due to the loss of degrees of freedom, these interactions were not included in the model.

$$Exp_{its} = \beta_{0s} + \beta_{1s} * \hat{h}_{its} + \sum_{m} \gamma_{m,s} * x_{m,its} + \phi_{ts} + \tau_{is} + \varepsilon_{s} \quad (9)$$

In the equations 7, 8 and 9 above, *Exp* represents a person's retirement expectations (either an age at which an individual plans to stop working or the probability with which an individual will be working full-time after the age of 62) of an individual i at time t with a health insurance type s, which is assumed to be a function of health $\hat{h}$, health insurance hi, their interaction $\hat{h}$*hi, other personal characteristics x_m, time indicator variables ϕ, and individual indicator variables τ. ε represents the error term. Because the probability of working full-time after the age of 62, *P(62)*, is a percentage bound between 0 and 1 with many values at the 0% and 100% level, I resort to using a probit model and a fixed effects logit model instead of a linear model.

6.3 Identification: Addressing Potential Biases

Justification Bias and Simultaneity As mentioned in section 3.4, a study of health's effects on retirement plans is made difficult by many possible biases, depending on the measures and the data used. First, I use panel data and look at marginal effects for those who are currently working. This may significantly improve my results in that if there is a state-dependent reporting, it is unlikely that for a working individual the bias would change in any direction over time in such short time periods. Similarly, the use of panel data and lagged health measure values may solve the issue of simultaneity. Additionally, the justification bias, even if it were present in self-reported health status, may not be a problem when examining the effects of a spouse's health on retirement plans because, although spouses may plan to retire together, they are interviewed separately and are less likely to report similar retirement ages or probabilities.

Endogeneity, Reverse Causality, and Classical Measurement Error Although the justification bias may not be present, there may still be sources of endogeneity that were so far unaccounted

for in previous studies. I focus on three main ones: leisure preferences, risk aversion, and parental health. In section 7, I include proxy variables for leisure preferences (number of close friends, average weekly hours), risk aversion (income risk aversion based on a series of questions), and parental health (whether mother and father are alive). Additionally, unlike other studies examining health and early retirement, I do not have a problem with possible reverse causality (early retirement affecting health and health affecting early retirement) because of the choice of the dependent variables[24]. Furthermore, because the errors-in-variables bias may be exaggerated in panel data estimation (Griliches and Hausman 1986), I use the latent variable model described by Bound et al. (1999, p. 186) and also used by Disney, Emmerson, and Wakefield (2006).

7 RESULTS

7.1 First Step Estimation

Table 8 displays the results of the ordered probit regression for years 2008 and 2010 from which the latent "health stock" values were obtained for the respondent. Table 9 displays results for spouses of the respondents (if present)[25]. The confusion matrix of the predictions for respondents can be seen in table 10[26]. Overall, 47.8% of the health status predictions matched the self-reported health status. The latent health status variable changes over time for 53.82% of the respondents. Figures 3 and 4 display the distributions of the variables estimated in the first step.

[24] Although, as was pointed out, the variables used in this study may be less informative than the actual age of retirement.

[25] Results for all years for respondents and spouses are available upon request.

[26] The confusion matrix for spouses' health is available upon request.

Table 8: Results of 2008 and 2010 ordered probit as a function of personal and health characteristics

	(1) 2008		(2) 2010	
	β	SE	β	SE
Activities				
Walking across a room	0.570**	(0.232)	−0.226	(0.417)
Walking a block	−0.130	(0.154)	−0.015	(0.209)
Walking several blocks	0.321***	(0.086)	0.445***	(0.123)
Getting in and out of bed	0.204	(0.221)	0.590*	(0.333)
Sitting for two hours	0.055	(0.071)	0.072	(0.092)
Getting up from a chair	0.161***	(0.056)	0.178**	(0.079)
Walking up a flight of stairs	−0.034	(0.105)	−0.167	(0.140)
Walking up several flights of stairs	0.462***	(0.058)	0.400***	(0.074)
Lifting 10 pounds	0.032	(0.086)	0.022	(0.113)
Picking up a small object (dime)	0.140	(0.153)	0.058	(0.168)
Health limits work	0.462***	(0.086)	0.432***	(0.111)
Medical care use				
Nights in hospital	0.022***	(0.008)	0.018	(0.012)
Nights in nursing home	−0.001	(0.001)	0.024***	(0.009)
Doctor visits	0.007***	(0.003)	0.008***	(0.002)
Illness occurence				
Has high blood pressure	0.371***	(0.047)	0.367***	(0.063)
Has diabetes	0.563***	(0.064)	0.514***	(0.082)
Has cancer	0.184**	(0.080)	0.351***	(0.098)
Has lung disease	0.184*	(0.110)	0.287**	(0.135)
Has heart problems	0.266***	(0.069)	0.215**	(0.087)
Has had a stroke	0.267	(0.182)	0.397*	(0.216)
Has psychological problems	0.155**	(0.062)	0.174**	(0.083)
Has arthritis	0.220***	(0.046)	0.113**	(0.057)
Total medical OOP exp. (in $1,000)	0.018***	(0.007)	0.009*	(0.005)
Demographic information				
Age at interview	−0.016***	(0.005)	−0.016**	(0.006)
Years of education	−0.052***	(0.011)	−0.048***	(0.014)
Total number of children	0.011	(0.012)	0.010	(0.015)
Female	−0.326***	(0.056)	−0.321***	(0.075)
Individual is black/African American	0.129*	(0.069)	0.146	(0.093)
Individial is Hispanic	0.072	(0.113)	0.245*	(0.146)
Individual is protestant	−0.147**	(0.065)	−0.088	(0.084)
Individual is catholic	−0.019	(0.072)	0.025	(0.093)
Individual is Jewish	−0.108	(0.156)	0.154	(0.172)
Marital status	−0.022	(0.052)	0.040	(0.068)
Mother is alive	−0.043	(0.043)	−0.087	(0.058)
Father is alive	−0.050	(0.052)	0.011	(0.069)
Observations	2928		1796	
Pseudo R^2	0.194		0.195	
Wald χ^2	1415.441		866.671	

Heteroskedasticity-robust standard errors in parentheses.
* p < 0.10, ** p < 0.51, *** p < 0.01.
Selected coefficients only, also included were prescription drugs, outpatient surgery, dentist services, special facilities use, activities of daily living score, negative feelings score, part-time employment, unemployed, veteran status, annual earnings, average weekly hours of work, occupation, industry, physical effort and stress at work, household wealth, pension, census region, unemployment rate.

Table 9: Results of 2008 and 2010 ordered probit as a function of spouse's personal and health characteristics

	(1) 2008		(2) 2010	
	β	SE	β	SE
Activities				
Walking across a room	0.113	(0.297)	−0.376	(0.685)
Walking a block	−0.159	(0.198)	0.055	(0.274)
Walking several blocks	0.544***	(0.121)	0.405***	(0.157)
Getting in and out of bed	0.013	(0.350)	0.736	(0.464)
Sitting for two hours	0.053	(0.105)	0.009	(0.118)
Getting up from a chair	0.218***	(0.076)	0.165	(0.102)
Walking up a flight of stairs	−0.157	(0.169)	−0.365*	(0.194)
Walking up several flights of stairs	0.439***	(0.081)	0.359***	(0.094)
Lifting 10 pounds	0.204	(0.126)	0.421***	(0.147)
Picking up a small object (dime)	−0.202	(0.214)	−0.066	(0.242)
Health limits work	0.266**	(0.123)	0.654***	(0.140)
Medical care use				
Nights in hospital	0.042**	(0.020)	0.016**	(0.007)
Nights in nursing home	−0.156***	(0.054)	−0.002	(0.003)
Doctor visits	0.014***	(0.005)	0.007*	(0.004)
Illness occurence				
Has high blood pressure	0.308***	(0.063)	0.484***	(0.079)
Has diabetes	0.478***	(0.085)	0.507***	(0.105)
Has cancer	0.100	(0.110)	0.130	(0.129)
Has lung disease	0.210	(0.160)	0.143	(0.181)
Has heart problems	0.380***	(0.090)	0.227**	(0.115)
Has had a stroke	−0.211	(0.215)	0.171	(0.271)
Has psychological problems	0.179**	(0.086)	0.263**	(0.109)
Has arthritis	0.289***	(0.061)	0.102	(0.075)
Total medical OOP exp. (in $1,000)	0.003	(0.005)	0.005	(0.007)
Demographic information				
Age at interview	−0.035***	(0.006)	−0.023***	(0.008)
Years of education	−0.039***	(0.013)	−0.052***	(0.015)
Female	−0.368***	(0.069)	−0.301***	(0.087)
Individual is black /African American	0.149	(0.109)	0.073	(0.135)
Individual is Hispanic	0.176	(0.119)	0.401***	(0.141)
Individual is protestant	−0.084	(0.093)	−0.047	(0.100)
Individual is catholic	−0.021	(0.103)	−0.131	(0.116)
Individual is Jewish	0.111	(0.201)	0.203	(0.240)
Mother is alive	−0.133**	(0.060)	−0.214***	(0.075)
Father is alive	−0.090	(0.068)	−0.060	(0.084)
Observations	1592		1086	
Pseudo R^2	0.184		0.195	
Wald χ^2	664.572		697.918	

Heteroskedasticity-robust standard errors in parentheses.
*p $<$ 0.10, ** p $<$ 0.51, *** p $<$ 0.01.
Selected coefficients only, also included were prescription drugs, outpatient surgery, dentist services, special facilities use, activities of daily living score, negative feelings score, employment status, veteran status, annual earnings, household wealth, pension, census region, unemployment rate.

Table 10: Confusion matrix of self-reported vs. predicted health status

		Self-reported health status				
		1: Excellent	2: Very good	3: Good	4: Fair	5: Poor
Predicted health status	1	3079	2111	531	24	1
	2	4902	9377	5306	657	23
	3	764	3115	5905	2540	248
	4	17	154	669	1213	314
	5	0	3	27	103	125
Mean of predicted probability of poor or fair health		0.032	0.066	0.155	0.349	0.587

Figure 3: Histograms of latent health status variables

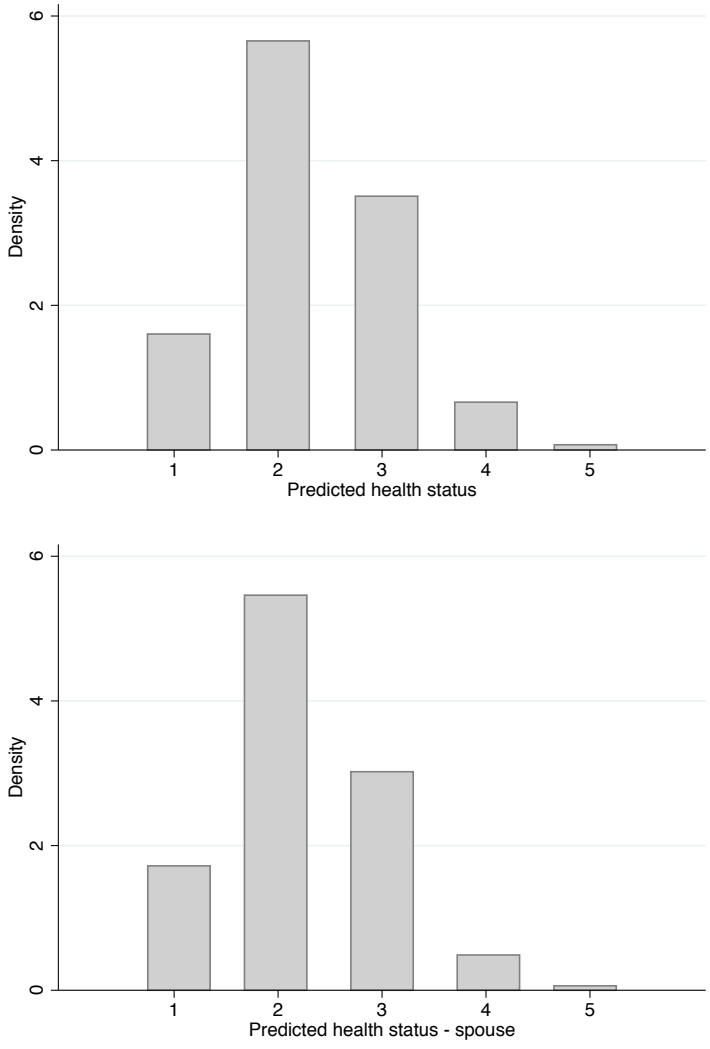
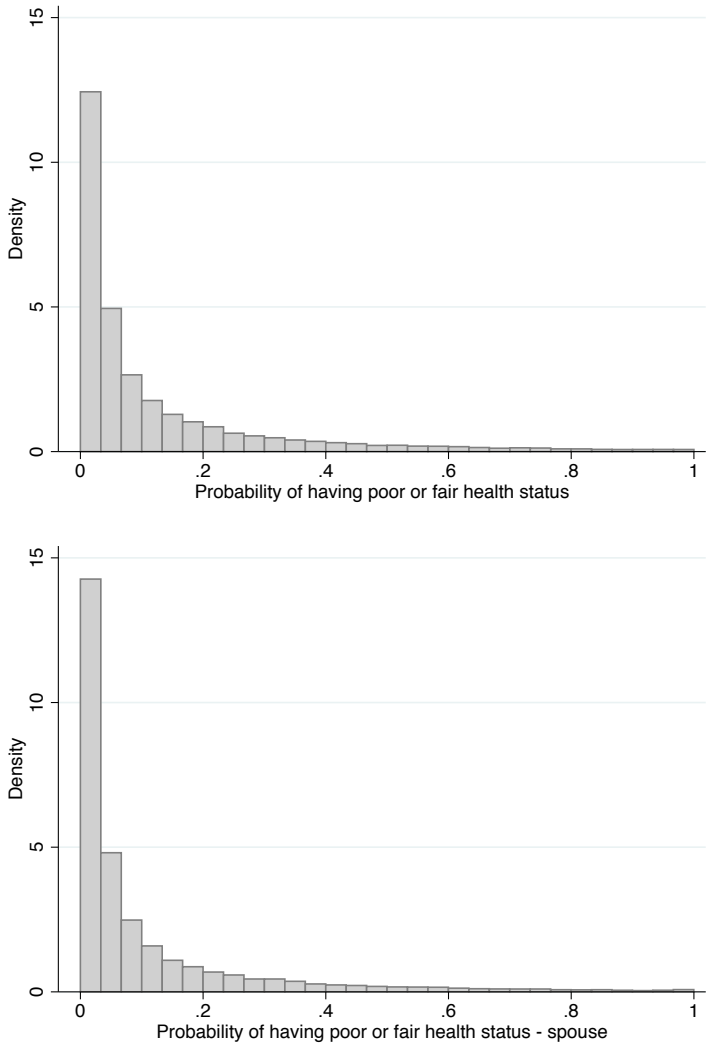


Figure 4: Histograms of the probability of having poor or fair health status



7.2 Pooled Model

Following the first step estimation of the latent health status, I first estimate the coefficients of interest on a pooled sample. This will allow us to see the type of relationship between health, health insurance, and retirement plans we may expect in the later fixed effects model. Section 7.2.1 discusses the results of a pooled sample analysis when the planned age of retirement is considered, while section 7.2.2 shows the results for the probability of working full-time past the age of 62. The fixed effects estimation results are discussed later in section 7.3.

7.2.1 Planned Retirement Age

Tables 11 and 12 display the regression results with planned retirement age as a dependent variable[27]. For both of the latent health variables, lower values mean overall better health status[28]. In all models, better health is associated with plans for a slightly later retirement; that is, if we disregard any interaction terms with early-retirement health insurance availability. The results in columns (1), (2), (3), (5), and (7) show that having health insurance available in case of an early retirement is generally associated with plans for an earlier retirement by approximately one year.

---

[27] The columns in both tables represent the different regression models that were run on the pooled sample of male household heads and of all respondents. In addition to the reported variables, the following control variables were included:

- Columns (1)-(7): age, age squared, gender, race (black), ethnicity (Hispanic), education, religion (protestant, catholic, Jewish), probability of living past the age of 75 (from Vital Statistics), marital status, number of children, cohort, whether mother/father are alive, employment status (part time, unemployed), veteran status, annual earnings (in $1,000), average weekly hours, work experience, experience squared, occupation, industry, physical effort and stress levels of a job, net household wealth (in $1,000), whether has pension, census division, census division unemployment rate, and region level health insurance policies.
- Column (4): whether covered by a union contract, job satisfaction, employer size, whether ever reported staying in current job because of health insurance, length of current marriage, how enjoyable is the time spent together with spouse, whether free time is spent with spouse, and whether was previously offered a retirement incentive.
- Column (6): Spouse's age, age squared, race(black), ethnicity (Hispanic), education, religion (protestant, catholic, Jewish), probability of living past the age of 75 (from Vital Statistics), whether mother/father are alive, employment status, veteran status, income, work experience, and whether has pension.

The standard errors were calculated using 200 bootstrap replications. I used 200 replications because according to Cameron and Trivedi (2010), running more than 200 replications of bootstrap does not bring any additional advantages.

[28] It was suggested that the latent health status be included as categories, not as one variable. While this approach may seem to be more interesting, it would not help answer the question of how an improved or worsening health affects retirement plans. For example, the coefficient on the 'good health status' category (assuming excellent was omitted), would show the effects of good health on retirement plans as compared to the effects of excellent health. However, this excludes the effects of any improvements in health. Therefore, in order to maintain flexibility of health status changes, the latent health status variable is not included in categories.

Table 11: Results of pooled linear model for R: males only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status	−0.257***	−0.265**	−0.115	−0.187	−0.734**	−0.695	
	(0.078)	(0.120)	(0.220)	(0.539)	(0.321)	(0.429)	
Predicted health squared					0.092		
					(0.057)		
Probability of having poor or fair health status							−1.095**
							(0.463)
Retiree HI	−1.360***						
	(0.256)						
Retiree HI * predicted health	0.316***						
	(0.108)						
Single — insured		−0.909***	−1.245**	0.110	−0.966***		−0.575***
		(0.318)	(0.582)	(1.218)	(0.320)		(0.137)
Single — insured * predicted health		0.209*	0.201	0.102	0.233*		
		(0.125)	(0.233)	(0.526)	(0.127)		
Single — insured * pr. of fair/poor health							1.196**
							(0.521)
Married — insured						−1.039	
						(0.986)	
Married — insured/uninsured						−2.663	
						(1.976)	
Married — uninsured/insured						0.501	
						(1.314)	
Married — insured * predicted health						0.530	
						(0.415)	
Married — insured/uninsured * predicted health						0.576	
						(0.744)	
Married — uninsured/insured * predicted health						0.452	
						(0.478)	
Income risk aversion			−0.070				
			(0.054)				
Individual has good friends in the neighborhood			−0.055				
			(0.151)				

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status - spouse						0.443	
						(0.323)	
Married — insured * predicted health (spouse)						−0.259	
						(0.316)	
Married — insured/uninsured * predicted health (spouse)						−0.073	
						(0.791)	
Married — uninsured/insured * predicted health (spouse)						−0.694	
						(0.431)	
Observations	8668	8668	2865	1261	8668	4552	8668
R^2	0.208	0.204	0.204	0.340	0.205	0.210	0.204

Bootstrap standard errors in parentheses. *$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

Table 12: Results of pooled linear model for R: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status	−0.116**	−0.161**	−0.048	−0.070	0.019	−0.188	
	(0.055)	(0.067)	(0.108)	(0.227)	(0.173)	(0.250)	
Predicted health squared					−0.036		
					(0.030)		
Probability of having poor or fair health status							−0.916***
							(0.266)
Retiree HI	−1.056***						
	(0.165)						
Retiree HI * predicted health	0.178***						
	(0.068)						
Single — insured		−0.771***	−0.905***	−0.391	−0.750***		−0.467***
		(0.183)	(0.285)	(0.597)	(0.186)		(0.080)
Single — insured * predicted health		0.195***	0.130	0.042	0.186***		
		(0.070)	(0.113)	(0.229)	(0.071)		
Single — insured * pr. of fair/poor health							1.152***
							(0.328)
Married — insured						−0.923	
						(0.707)	
Married — insured/uninsured						−3.076**	
						(1.534)	
Married — uninsured/insured						−0.162	
						(0.867)	
Married — insured * predicted health						0.178	
						(0.262)	
Married — insured/uninsured * predicted health						0.050	
						(0.564)	
Married — uninsured/insured * predicted health						0.136	
						(0.292)	
Income risk aversion			−0.046				
			(0.035)				
Individual has good friends in the neighborhood			−0.168*				
			(0.090)				

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status - spouse						0.162	
						(0.192)	
Married — insured * predicted health (spouse)						0.074	
						(0.194)	
Married — insured/uninsured * predicted health (spouse)						0.513	
						(0.460)	
Married — uninsured/insured * predicted health (spouse)						−0.149	
						(0.221)	
Observations	21818	21818	9660	2840	21818	10064	21818
R^2	0.174	0.170	0.159	0.249	0.170	0.187	0.171

Bootstrap standard errors in parentheses. *$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

Additionally, the results in columns (1), (2), (5) and (7) offer an insight into how health insurance availability and health may interact in the realm of retirement planning. It seems that for those without an early-retirement health insurance option, better health will lead to a later retirement, while the relationship between latent health and retirement plans is diminished in case of those who have an early-retirement health insurance available. This may suggest that those who have health insurance available to them in case of an early retirement plan to retire earlier than others regardless of their health. That is, health as such does not matter very much to these individuals when making plans for retirement as they will be insured regardless of whether or not they are employed until the age of 65.

Tables 11 and 12 also include the results when controlling for risk aversion, leisure preferences, health squared, and spouse's health and characteristics. When risk aversion and leisure preferences were included in the model, the effects of health on retirement plans became insignificant, while the magnitude of the effect of health insurance increased by a small amount. Although this effect could be due the two control variables, it is entirely because of the smaller sample as risk aversion and the information on the respondent's friends were largely missing[29].

When health squared was added in the model, the effect of predicted health status became much larger in magnitude in case of males, but became small and insignificant for all respondents while the coefficient on health insurance did not change much[30]. Finally, when a spouse's health is added in the model and the health insurance options after early retirement are divided into four groups (as described in appendix C), while almost all coefficients keep the expected signs, most are

[29]A separate regression was run where the sample was restricted to only those individuals, for whom risk aversion and leisure preferences were available, but with risk aversion and leisure preferences excluded. The results were indistinguishable from the model with risk aversion and leisure preferences included.

[30]However, the test of joint significance of health and health squared failed to reject the joint insignificance of these two variables.

statistically insignificant. It is important to note that the comparison group here is the case in which both the respondent and his or her spouse remain uninsured after the respondent retires before being eligible for Medicare. In the case of all respondents, it seems that if a respondent's early retirement would leave the respondent with insurance, while leaving the spouse without insurance, the respondent would tend to retire all of three years earlier than if the couple were to be uninsured. Additionally, spouse's health status does not seem to be significantly affecting the respondent's retirement plans.

7.2.2 Probability of Working Full-time Past the Age of 62

The results for a probit model applied to a pooled sample of males and all respondents can be viewed in tables 13 and 14 respectively. The tables are structured identically to the tables in section 7.2.1[31]. In these tables, the coefficients are expected to have the same signs as those in tables 13 and 14[32]. However, it is good to keep in mind that these coefficients represent the average marginal effects. Overall, effects do have the same direction as in the case of planned retirement age.

[31] For variables included in each regression, please, see footnote 27. The standard errors were calculated using 200 bootstrap replications.

[32] The average marginal effects are displayed instead of the actual probit coefficients. Average marginal effects were estimated following each probit regression and thus are displayed in all probit model results instead of the actual probit coefficients, which are less meaningful.

Table 13: Results of pooled probit model for *P62*: males only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status	−0.021***	−0.021**	−0.031**	−0.043	0.011	−0.064***	
	(0.006)	(0.008)	(0.013)	(0.063)	(0.020)	(0.018)	
Predicted health squared					−0.006*		
					(0.003)		
Probability of having poor or fair health status							−0.095***
							(0.031)
Retiree	−0.069***						
	(0.017)						
Retiree HI * predicted health	0.003						
	(0.007)						
Single — insured		−0.041*	−0.072**	−0.110	−0.038*		−0.032***
		(0.022)	(0.036)	(0.155)	(0.022)		(0.010)
Single — insured * predicted health		0.003	0.014	0.020	0.001		
		(0.009)	(0.014)	(0.061)	(0.009)		
Single — insured * pr. of fair/poor health							0.044
							(0.035)
Married — insured						−0.134***	
						(0.050)	
Married — insured/uninsured						−0.289*	
						(0.167)	
Married — uninsured/insured						−0.174***	
						(0.056)	
Married — insured * predicted health						0.034*	
						(0.019)	
Married — insured/uninsured * predicted health						−0.006	
						(0.050)	
Married — uninsured/insured * predicted health						0.065***	
						(0.023)	
Income risk aversion			−0.009***				
			(0.003)				
Individual has good friends in the neighborhood			0.016*				
			(0.009)				

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status - spouse						−0.004	
						(0.016)	
Married — insured * predicted health (spouse)						0.013	
						(0.015)	
Married — insured/uninsured * predicted health (spouse)						0.098**	
						(0.050)	
Married — uninsured/insured * predicted health (spouse)						0.005	
						(0.021)	
Observations	14799	14799	5473	1480	14799	7848	14799
Pseudo R^2	0.094	0.088	0.117	0.113	0.088	0.108	

Average marginal effects displayed (instead of MEM). Bootstrap standard errors in parentheses. *$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

Table 14: Results of pooled probit model for *P62*: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status	−0.011***	−0.012***	−0.011	0.014	0.042***	−0.022	
	(0.004)	(0.004)	(0.007)	(0.022)	(0.012)	(0.015)	
Predicted health squared					−0.011***		
					(0.002)		
Probability of having poor or fair health status							−0.074***
							(0.017)
Retiree	−0.058***						
	(0.012)						
Retiree HI * predicted health	0.001						
	(0.005)						
Single — insured		−0.036***	−0.045**	0.011	−0.029**		−0.029***
		(0.012)	(0.022)	(0.057)	(0.012)		(0.005)
Single — insured * predicted health		0.004	0.003	−0.017	0.001		
		(0.005)	(0.009)	(0.023)	(0.005)		
Single — insured * pr. of fair/poor health							0.015
							(0.022)
Married — insured						−0.075*	
						(0.042)	
Married — insured/uninsured						−0.260**	
						(0.119)	
Married — uninsured/insured						−0.092**	
						(0.041)	
Married — insured * predicted health						0.005	
						(0.015)	
Married — insured/uninsured * predicted health						−0.045	
						(0.046)	
Married — uninsured/insured * predicted health						0.019	
						(0.016)	
Income risk aversion			−0.006***				
			(0.002)				
Individual has good friends in the neighborhood			0.005				
			(0.007)				

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status - spouse						0.008	
						(0.012)	
Married — insured * predicted health (spouse)						0.015	
						(0.012)	
Married — insured/uninsured * predicted health (spouse)						0.120***	
						(0.039)	
Married — uninsured/insured * predicted health (spouse)						0.008	
						(0.012)	
Observations	35613	35613	16046	3555	35613	16633	35613
Pseudo R^2	0.082	0.079	0.092	0.086	0.080	0.098	0.080

Bootstrap standard errors in parentheses. *$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

Average marginal effects displayed (instead of MEM) due to many binary demographic variables in the model.

When it comes to health insurance, those who have an early-retirement health insurance available are less likely to work full time after the age of 62 by 3.2 to 7.2 percentage points when looking at males only, while the coefficients are slightly smaller in magnitude when looking at the entire sample. The way health status affects the probability of working full time after the age of 62 seems to depend on health insurance; although, the interaction coefficients are not significant even at the 10% level. Thus, we may only conclude that those of worse health will have a lower probability of working full-time after the age of 62. If one's health declines by one point (moving from very good to good for example), the individual will be less likely to keep working by 2.1 to 6.4 percentage points, depending on the model specification.

Model in column (3) in both tables includes measures of risk aversion and leisure (represented by an indicator of having friends in one's neighborhood). Here, lower value of the risk aversion variable means lower risk aversion of an individual. Thus, it seems that being less risk averse is associated with a higher probability of working past the age of 62. Similarly, those with friends in their neighborhood (which may represent higher leisure preferences) tend to be more likely to work after the age of 62. The results are contrary to what one would expect, but they are of a very small in magnitude in both tables.

The results for a model that contains health squared are displayed in column (5). These results seem to confirm an earlier theory that, because of the way health may affect household assets, individuals may choose to retire later due to a worsening health, but only until a certain point. As their health becomes fair or poor, individuals may actually begin to plan an earlier retirement and the probability of working full time after the age of 62 decreases.

Finally, column (6) displays the results for married individuals when their spouses' health and health insurance are considered in addition to their own. Compared to an uninsured couple, all others have a lower probability of working full time after the age of 62 - between 7.5 and 26 percent-

age points for all respondents and between 13.4 and 28.9 percentage points for male respondents. A few of the interaction variables also suggest that if a couple is less likely to lose health insurance upon early retirement, the effect of health on the probability of working full time after the age of 62 diminishes.

7.3 Fixed Effects Estimation

Once the impact of health and health insurance was examined in detail in the pooled model, we now turn to a panel data structure for a more insightful analysis. Here, both theory and statistical testing have confirmed that the fixed effects (within) estimator be used.

7.3.1 Planned Retirement Age

Tables 15 and 16 display the results of a fixed effects regression model[33]. The problem with this type of a model is that over time, health status and insurance change very little for individual respondents; thus, many coefficients are statistically insignificant because we cannot identify the effects as well as in a pooled model.

[33]The individual model specifications are similar to what is described in footnote 27. However, column (6) includes only the basic control variables (meaning no spouse control variables are used), while column (7) includes the same control variables as column (6) did in footnote 27. Column (8) contains only the basic control variables. The standard errors were calculated using 200 bootstrap replications. There were not enough observations to compute bootstrap standard errors in columns (3), (4) and (7). Huber-White standard errors were calculated instead.

Table 15: Results of fixed effects linear model for *R*: males only

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted health status	−0.110	−0.020	0.336	−0.070	−0.047	0.085	−0.041	
	(0.103)	(0.192)	(0.408)	(0.570)	(0.331)	(0.370)	(0.488)	
Predicted health squared					0.017	0.013		
					(0.053)	(0.063)		
Lagged health status						−0.047		
						(0.067)		
Probability of having poor or fair health status								0.033
								(0.854)
Retiree HI	−0.204							
	(0.271)							
Retiree HI * predicted health	0.196*							
	(0.112)							
Single — insured		0.304	0.134	1.511	0.354	0.571		0.389**
		(0.441)	(1.062)	(1.371)	(0.464)	(0.514)		(0.175)
Single — insured * predicted health		0.016	−0.015	0.037	0.002	−0.125		
		(0.178)	(0.415)	(0.587)	(0.193)	(0.197)		
Single — insured * pr. of fair/poor health								−0.261
								(0.864)
Married — insured							−0.363	
							(1.297)	
Married — insured/uninsured							2.035	
							(1.881)	
Married — uninsured/insured							−0.283	
							(1.488)	
Married — insured * predicted health							0.101	
							(0.490)	
Married — insured/uninsured * predicted health							−0.981	
							(0.701)	
Married — uninsured/insured * predicted health							−0.008	
							(0.562)	
Income risk aversion			−0.058					
			(0.064)					
Individual has good friends in the neighborhood			0.691***					
			(0.230)					

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted health status - spouse							0.200	
							(0.497)	
Married — insured * predicted health (spouse)							0.064	
							(0.492)	
Married — insured/uninsured * predicted health (spouse)							0.250	
							(0.692)	
Married — uninsured/insured * predicted health (spouse)							−0.080	
							(0.543)	
Observations	6815	6815	2466	892	7094	6142	3862	7094
R^2	0.090	0.089	0.130	0.423	0.067	0.074	0.116	0.067

Bootstrap standard errors in parentheses. Columns (3), (4) and (7): Huber-White standard errors in parentheses. *$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

Table 16: Results of fixed effects linear model for R: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted health status	−0.078	−0.084	0.082	−0.633***	−0.065	−0.019	−0.168	
	(0.067)	(0.078)	(0.152)	(0.148)	(0.205)	(0.236)	(0.254)	
Predicted health squared					−0.004	−0.007		
					(0.040)	(0.046)		
Lagged health status						0.008		
						(0.050)		
Probability of having poor or fair health status								−0.734**
								(0.344)
Retiree HI	−0.012							
	(0.189)							
Retiree HI * predicted health	0.038							
	(0.078)							
Single — insured		0.078	0.159	−0.103	0.080	0.134		0.126
		(0.212)	(0.417)	(0.428)	(0.213)	(0.223)		(0.097)
Single — insured * predicted health		0.033	−0.047	0.384***	0.033	0.004		
		(0.082)	(0.175)	(0.094)	(0.083)	(0.090)		
Single — insured * pr. of fair/poor health								0.250
								(0.367)
Married — insured							−0.614	
							(0.682)	
Married — insured/uninsured							0.899	
							(1.440)	
Married — uninsured/insured							−0.851	
							(0.732)	
Married — insured * predicted health							0.095	
							(0.216)	
Married — insured/uninsured * predicted health							−0.250	
							(0.357)	
Married — uninsured/insured * predicted health							0.112	
							(0.321)	
Income risk aversion			−0.122***					
			(0.040)					
Individual has good friends in the neighborhood			−0.035					
			(0.154)					

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted health status - spouse							−0.164	
							(0.307)	
Married — insured * predicted health (spouse)							0.276	
							(0.332)	
Married — insured/uninsured * predicted health (spouse)							−0.377	
							(0.800)	
Married — uninsured/insured * predicted health (spouse)							0.250	
							(0.354)	
Observations	18141	18141	7692	2175	18141	16098	8348	18141
R^2	0.095	0.095	0.112	0.255	0.095	0.099	0.082	0.095

Bootstrap standard errors in parentheses. Columns (3), (4) and (7): Huber-White standard errors in parentheses. *$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

While the coefficients on health status and remain generally of a similar magnitude and direction, the only statistically significant coefficient can be found when latent health is considered in table 16 in column (4). Here, it seems that a change in health by one category (from excellent to very good, or good to fair for example) is associated by a shift in retirement plans of approximately 7.5 months earlier for those with no early-retirement health insurance. For those individuals with health insurance security, this effect diminishes to about three months, which demonstrates the fact that health matters less in retirement planning when individuals do not lose health insurance if they retire early.

Due to the fact that individual respondents seldom change health insurance over time and because the characteristics across different health insurance groups are not very different, tables 17 and 18 display the fixed effect model results when only those with a particular health insurance are included[34]. When it comes to a respondent's own health and the age at which he/she plans to retire, it seems that health matters more to those without early-retirement health insurance available. This can be seen especially in columns (3) and (4) of table 17, and columns (1) and (2), and (3) and (4) of table 18. Although the effects are more modest in case of the latent health status, they are still of a similar magnitude and the same direction.

In case of a spouse's health (columns (5)-(6) of both tables), for couples who would remain insured in case of an early retirement, as a spouse's health worsens, the respondent's retirement plans change to a later retirement, while the opposite is true for those who would remain or become uninsured. The latter result is driven completely by the males in the sample, whose spouses are very sick – by the need for care giving. The former result is then driven mostly by those whose spouses

[34]Including only those with a particular health insurance type results in a smaller sample size. In all columns, the basic control variables were included, while columns (5)-(7) also include spouse characteristics. See footnote 27 for a comprehensive listing of all control variables. The standard errors were calculated using 200 bootstrap replications, except for columns (5)- (7) where Huber-White standard errors were used

Table 17: Results of fixed effects linear model for R, latent health status, by health insurance type: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Predicted health status	−0.030	−0.087	−0.044	−0.219*	−0.100	0.046	−0.501
	(0.086)	(0.081)	(0.071)	(0.116)	(0.102)	(0.209)	(0.691)
Predicted health status - spouse					0.219**	−0.298	−1.637*
					(0.101)	(0.226)	(0.937)
Annual earnings (in $1,000)	−0.001	−0.004***	−0.002*	−0.005***	−0.001	−0.003*	0.008
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.021)
Net household wealth (in $1,000)	0.000	0.000***	0.000	0.000*	0.000	0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Observations	7175	10966	11668	6473	5743	1953	544
R^2	0.117	0.104	0.101	0.124	0.086	0.177	0.372

Columns (1)-(4): Bootstrap standard errors in parentheses. Columns (5)-(7): Huber-White standard errors in parentheses. * $p <$ *0.10,* ** $p <$ *0.51,* *** $p <$ *0.01*

The columns represent the following health insurance groups: (1) retiree health insurance; (2) no retiree health insurance; (3) single – insured; (4) single – uninsured; (5) married – insured; (6) married – uninsured/insured; and (7) married –uninsured. Please, see appendix C for a definition of these groups.

Table 18: Results of fixed effects linear model for R, probability of poor or fair health, by health insurance type: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Probability of having poor or fair health status	−0.485	−0.758**	−0.475	−1.467***	0.034	−0.301	−3.601
	(0.447)	(0.361)	(0.398)	(0.520)	(0.554)	(1.114)	(4.598)
Probability of having poor or fair health status - spouse					0.203	−1.996	−9.074**
					(0.565)	(1.484)	(4.343)
Annual earnings (in $1,000)	−0.001	−0.004***	−0.002*	−0.005***	−0.001	−0.003*	0.009
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.022)
Net household wealth (in $1,000)	0.000	0.000***	0.000	0.000	0.000	0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Observations	7175	10966	11668	6473	5743	1953	544
R^2	0.118	0.104	0.101	0.125	0.084	0.179	0.378

Columns (1)-(4): Bootstrap standard errors in parentheses. Columns (5)-(7): Huber-White standard errors in parentheses. * $p <$ *0.10,* ** $p <$ *0.51,* *** $p <$ *0.01*

The columns represent the following health insurance groups: (1) retiree health insurance; (2) no retiree health insurance; (3) single – insured; (4) single – uninsured; (5) married – insured; (6) married – uninsured/insured; and (7) married –uninsured. Please, see appendix C for a definition of these groups.

are relatively healthy and where a change of health from excellent to very good or very good to good does not require care giving but may require the spouse to leave work, leaving the respondent as a sole income provider[35].

7.3.2 Probability of Working Full-time Past the Age of 62

Tables 19 and 20 display the results of fixed effects estimation when the probability of working full-time after the age of 62 becomes a dependent variable. In this case, it was not always possible to estimate a fixed effects model. While with the age at which one plans to stop working (or the planned retirement age) the Hausman test did suggest that fixed effects are the appropriate model, this was not so in the case of the probability of working full-time past the age of 62[36]. Thus, even though in theory, fixed effects seem to be the correct model to be used, the random effects logit model was used where the fixed effects logit model could not be estimated[37].

---

[35]These conclusions are based on additional regressions, the results of which are not included in this study due to space constraints. The results are available from the author upon request.

[36]The Hausman test compared a fixed effects model with a random effects model. Here, the Hausman test's χ^2 statistic was high and the test failed to reject the null at the 5% significance level.

[37]The individual model specifications are similar to what is described in footnote 27. However, column (6) includes only the basic control variables (meaning no spouse control variables are used), while column (7) includes the same control variables as column (6) did in footnote 27. Column (8) contains only the basic control variables. Random effects model (instead of a fixed effects model) was estimated in columns (3) and (7) of table 19 and in columns (2) and (4) of table 20. Although the random effects model does not rid the model of endogeneity, it was used here after the fixed effects model could not be estimated.

Table 19: Results of fixed effects logit model for *P62*: males only

	(1)	(2)	(3)	(5)	(6)	(7)	(8)
Predicted health status	−0.000	−0.000	−0.030***	−0.001	0.013	−0.043**	
	(0.002)	(0.003)	(0.010)	(0.005)	(0.044)	(0.019)	
Predicted health squared				0.000	−0.003		
				(0.000)	(0.008)		
Lagged health status					0.009		
					(0.010)		
Probability of having poor or fair health status							−0.001
							(0.009)
Retiree HI	0.000						
	(0.002)						
Retiree HI * predicted health	−0.001						
	(0.005)						
Single — insured		0.000	−0.073***	0.000	0.027		0.000
		(0.003)	(0.028)	(0.003)	(0.052)		(0.002)
Single — insured * predicted health		−0.000	0.021*	−0.000	−0.012		
		(0.001)	(0.011)	(0.001)	(0.021)		
Single — insured * pr. of fair/poor health							−0.001
							(0.006)
Married — insured						−0.110**	
						(0.056)	
Married — insured/uninsured						−0.284**	
						(0.111)	
Married — uninsured/insured						−0.131**	
						(0.065)	
Married — insured * predicted health						0.022	
						(0.020)	
Married — insured/uninsured * predicted health						0.028	
						(0.041)	
Married — uninsured/insured * predicted health						0.046**	
						(0.023)	
Income risk aversion			−0.006**				
			(0.003)				

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(5)	(6)	(7)	(8)
Individual has good friends in the neighborhood			0.016**				
			(0.007)				
Predicted health status - spouse						−0.009	
						(0.017)	
Married — insured * predicted health (spouse)						0.015	
						(0.018)	
Married — insured/uninsured * predicted health (spouse)						0.062	
						(0.039)	
Married — uninsured/insured * predicted health (spouse)						0.002	
						(0.020)	
Observations	3636	3636	5139	3636	2953	7336	3636
Pseudo R^2	0.032	0.029		0.029	0.038		0.028
Wald χ^2 (d.f.)	84.710(61)	76.699(61)	223.613(71)	76.719(62)	82.008(61)	319.348(95)	76.158(61)

Average marginal effects. Standard errors in parentheses.
*$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

Table 20: Results of fixed effects logit model for *P62*: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted health status	−0.002***	−0.012**	−0.006*	0.008	0.000	0.005	−0.003	
	(0.004)	(0.005)	(3.285)	(0.024)	(0.002)	(0.016)	(1.448)	
Predicted health squared					−0.000	−0.003		
					(0.001)	(0.008)		
Lagged health status						−0.001		
						(0.003)		
Probability of having poor or fair health status								−0.007***
								(0.018)
Retiree	−0.001							
	(0.003)							
Retiree HI * predicted health	−0.000							
	(0.001)							
Single — insured		−0.028**	−0.005	−0.009	−0.001	−0.003		−0.001
		(0.013)	(2.706)	(0.062)	(0.003)	(0.013)		(0.002)
Single — insured * predicted health		0.002	0.002	−0.010	0.000	−0.001		
		(0.005)	(1.185)	(0.026)	(0.001)	(0.005)		
Single — insured * pr. of fair/poor health								−0.000
								(0.002)
Married — insured							0.005	
							(2.576)	
Married — insured/uninsured							−0.051	
							(26.989)	
Married — uninsured/insured							0.008	
							(4.473)	
Married — insured * predicted health							−0.003	
							(1.473)	
Married — insured/uninsured * predicted health							−0.002	
							(1.009)	
Married — uninsured/insured * predicted health							−0.004	
							(1.906)	
Income risk aversion			−0.000					
			(0.201)					
Individual has good friends in the neighborhood			0.001					
			(0.726)					

Continued on next page

Continued from previous page

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted health status - spouse							0.004	
							(2.313)	
Married — insured * predicted health (spouse)							−0.004	
							(2.335)	
Married — insured/uninsured * predicted health (spouse)							0.012	
							(6.237)	
Married — uninsured/insured * predicted health (spouse)							−0.006	
							(3.253)	
Observations	11160	33353	2668	3122	11160	9334	4399	11160
Pseudo R^2	0.025		0.053		0.024	0.026	0.038	0.024
Wald χ^2 (d.f.)	209.026(61)	1251.437(71)	102.284(61)	186.098(82)	204.865(62)	182.453(62)	126.313(83)	205.329(61)

Average marginal effects. Standard errors in parentheses.
*$p < 0.10$, ** $p < 0.51$, *** $p < 0.01$

When heath status is considered, the direction of the relationship remains the same, with a larger magnitude. As the results in other tables suggest, the worsening of health by one point (moving from one category to another) may lead to a 0.2 to 4.3 percentage points lower probability of working full-time after the age of 62. This is very similar to what was estimated in the pooled model. However, most of the health and health insurance interaction coefficients are very close to zero and thus it is difficult to conclude whether the effect of health on retirement plans varies with health insurance. This is mostly due to the fact that health insurance varies very little over time. Finally, having early-retirement health insurance available lowers the probability of working full-time after the age of 62 by 2.8 to 7.3 percentage points. This is very similar to what was estimated earlier in the pooled model.

As stated in section 7.3.1, the low variation in health status and health insurance over time are the main reasons why so many of the coefficients in the fixed effects models are statistically insignificant. Therefore, tables 21 and 22 display the results of the panel models when only those individuals with a particular health insurance are included[38]. Respondent's own health seems to have a larger effect on the probability of working full time after the age of 62 when the respondent would retain insurance in case of an early retirement in columns (1) and (2), while the relationship is reversed in columns (3) and (4). The results in column (5) are very similar in nature to those in tables 17 and 18 with the same interpretation. Overall, it seems that for all types of early-retirement health insurance options, the respondent will tend to retire early if his or her health declines.

[38]Including only those with a particular health insurance type results in smaller sample sizes. In all columns, the basic control variables were included, while columns (5) and (6) include spouse characteristics in addition to the basic respondent's characteristics. See footnote 27 for a comprehensive listing of all control variables. Random effects model was estimated in columns (1), (2) and (4)-(6) of both tables.

Table 21: Results of fixed effects logit model for *P62*, latent health status, by health insurance type: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)
Predicted health status	-0.016^{**}	-0.008^{**}	-0.000^{*}	-0.012^{**}	-0.016^{**}	-0.009
	(0.006)	(0.004)	(0.014)	(0.005)	(0.006)	(0.013)
Predicted health status - spouse					0.016^{***}	0.005
					(0.006)	(0.012)
Annual earnings (in $1,000)	-0.000	0.000	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Net household wealth (in $1,000)	-0.000^{***}	-0.000^{**}	-0.000	-0.000^{*}	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	11539	21814	5678	13810	9868	1506
Wald χ^2	448.127	962.194	126.584	709.891	428.403	68.715
D.f.	69	69	59	69	86	85

Average marginal effects. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.51$, *** $p < 0.01$
The columns represent the following health insurance groups: (1) retiree health insurance; (2) no retiree health insurance; (3) single – insured; (4) single – uninsured; (5) married – insured; and (6) married –uninsured. Please, see appendix C for a definition of these groups.

Table 22: Results of fixed effects logit model for *P62*, probability of poor or fair health, by health insurance type: all respondents

	(1)	(2)	(3)	(4)	(5)	(6)
Probability of having poor or fair health status	-0.085^{***}	-0.053^{***}	-0.000^{**}	-0.073^{***}	-0.096^{***}	-0.037
	(0.029)	(0.015)	(0.059)	(0.019)	(0.030)	(0.053)
Probability of having poor or fair health status - spouse					0.062^{**}	0.005
					(0.030)	(0.047)
Annual earnings (in $1,000)	-0.000	0.000	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Net household wealth (in $1,000)	-0.000^{***}	-0.000^{**}	-0.000	-0.000^{*}	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	11539	21814	5678	13810	9868	1506
Wald χ^2	449.060	967.113	128.406	712.461	428.964	69.385
D.f.	69	69	59	69	86	85

Average marginal effects. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.51$, *** $p < 0.01$
The columns represent the following health insurance groups: (1) retiree health insurance; (2) no retiree health insurance; (3) single – insured; (4) single – uninsured; (5) married – insured; and (6) married –uninsured. Please, see appendix C for a definition of these groups.

8 CONCLUSION

This study has described the effects of health on retirement of individuals and couples and how these effects vary with the different sources of health insurance. I used the biennial Health and Retirement Study data ranging from years 1992 to 2010 with two unique dependent variables: the planned retirement age and the probability of working full-time past the age of 62. This allowed me to conduct a longitudinal study and thus avoid some identification problems found in the previous literature.

Overall, the results suggest that the effects of health on retirement plans do indeed vary with health insurance. This was tested both cross-sectionally and longitudinally. It was estimated that those who lack any early-retirement health insurance options are affected more than others by health in their retirement planning. Generally, these individuals are more likely to retire earlier only if their health worsens, while those who retain health insurance upon an early retirement tend to retire earlier no matter what. Surprisingly, spouse's health tends not to affect retirement plans very much; although, as some fixed effects results suggest, a very sick spouse may result in an individual's early retirement (or the plans therefor). The results also document that the relationship of health and retirement plans may not be linear, but quadratic, where individuals are more likely to retire early only after health worsens to a certain level.

These results are very important in that they shed light upon a long line of conflicting research results. They give an explanation as to why some earlier studies may have ended with opposing results and may serve as a great tool to policy makers. Based on this study, one may, for example, predict some of the consequences of the Patient Protection and Affordable Care Act, under which all individuals will gain an equal access to health insurance.

The main limitations of the study include the fact that the two dependent variables are

less precise (compared to an actual age of retirement) in reflecting individuals' retirement choices. Thus, the results may lose precision as a consequence of the dependent variable choice. Additionally, small response rates for the planned age of retirement resulted in small sample size in the fixed effects model. Finally, the low variation over time in the variables of interest lead to many insignificant results. A great extension of this study would be an examination of why individuals' retirement plans differ from the actual retirement and of the main determinants of this phenomenon.

WORKS CITED

Baltagi, Badi H. 2008. *Econometric analysis of panel data.* 4th. Chichester, West Sussex, UK: John Wiley & Sons Ltd.

Blau, David M., and Donna B. Gilleskie. 2001. Retiree health insurance and the labor force behavior of older men in the 1990s. *The Review of Economics and Statistics* 83 (1): 64–80.

———. 2006. Health insurance and retirement of married couples. *Journal of Applied Econometrics* 21 (7): 935–953.

Bound, John. 1991. Self-reported versus objective measures of health in retirement models. *Journal of Human Resources* 26 (1): 106–138.

Bound, John, et al. 1999. The dynamic effects of health on the labor force transitions of older workers. *Labour Economics* 6 (2): 179–202.

Bureau of Labor Statistics. 2011. *Local area unemployment statistics.*

Cameron, A. Colin, and Pravin K. Trivedi. 2010. *Microeconometrics using stata.* 2nd. Stata Press.

Centers for Disease Control and Prevention. 2012a. *Asthma and chronic obstructive pulmonary disease, 2008-2010.*

———. 2012b. *Chronic conditions, ages 18+, 2007-2009.*

———. 2012c. *Health status, respondent reported, 2008-2010.*

———. 2012d. *Hypertension, ages 20+, 2008-2010.*

Datta Gupta, Nabanita, and Mona Larsen. 2010. The impact of health on individual retirement plans: self-reported versus diagnostic measures. *Health Economics* 19 (7): 792–813.

Disney, Richard, Carl Emmerson, and Matthew Wakefield. 2006. Ill health and retirement in britain: a panel data-based analysis. *Journal of Health Economics* 25 (4): 621–649.

Dwyer, Debra Sabatini, and Olivia S. Mitchell. 1999. Health problems as determinants of retirement: are self-rated measures endogenous? *Journal of Health Economics* 18 (2): 173 –193. http://www.sciencedirect.com/science/article/pii/S0167629698000344.

Federal Reserve Economic Data. 1992-2010. *Real gross domestic product.*

French, Eric, and John Bailey Jones. 2011. The effects of health insurance and self-insurance on retirement behavior. *Econometrica* 79 (3): 693–732.

Green, Carole A. 2006. The unexpected impact of health on the labor supply of the oldest americans. *Journal of Labor Research* 27 (3): 361–379.

Griliches, Zvi, and Jerry A. Hausman. 1986. Errors in variables in panel data. *Journal of Econometrics* 31 (1): 93 –118. http://www.sciencedirect.com/science/article/pii/0304407686900588.

Gruber, Jonathan. 2009. *How today's health care reform can create tomorrow's entrepreneurs.*

Gruber, Jonathan, and Brigitte C. Madrian. 1995. Health-insurance availability and the retirement decision. *The American Economic Review* 85 (4): 938–948.

———. 1996. *Health insurance and early retirement: evidence from the availability of continuation coverage,* ed. David A. Wise, 115–143. National Bureau of Economic Research Project Report series. Chicago / London: University of Chicago Press.

———. 2002. Health insurance, labor supply, and job mobility: a critical review of the literature. *NBER Working Paper Series.*

Health and Retirement Study. 1992-2010a. Public use dataset (RAND HRS Fat files).

Health and Retirement Study. 1992-2010b. Public use dataset (RAND HSR data file, version L).

———. 2011. Sample sizes and response rates.

Hurd, Michael D. 1990. *The joint retirement decision of husbands and wives,* ed. David A. Wise, 231–254. A National Bureau of Economic Research Project Report Chicago / London: University of Chicago Press.

Jacobson, Gretchen, Karyn Schwartz, and Tricia Neuman. 2009. Health insurance coverage for older adults: implications of a medicare buy-in.

Jones, Andrew M., Nigel Rice, and Jennifer Roberts. 2010. Sick of work or too sick to work? evidence on self-reported health shocks and early retirement from the bhps. *Economic Modelling* 27 (4): 866–880.

Kapur, Kanika, and Jeannette Rogowski. 2007. The role of health insurance in joint retirement among married couples. *Industrial and Labor Relations Review* 60 (3): 397–407.

Lindeboom, Maarten, and Marcel Kerkhofs. 2009. Health and work of the elderly: subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics* 24 (6): 1024–1046.

Madrian, Brigitte C., and Nancy Dean Beaulieu. 1998. *Does medicare eligibility affect retirement?* Ed. David A. Wise, 109–128. NBER Project Report series. Chicago / London: University of Chicago Press.

Madrian, Brigitte C., Gary Burtless, and Jonathan Gruber. 1994. The effect of health insurance on retirement. *Brookings Papers on Economic Activity* 1994 (1): 181–252.

McGarry, Kathleen. 2004. Health and retirement: do changes in health affect retirement expectations? *Journal of Human Resources* 39 (3): 624–648.

Miah, M. Solaiman, and Virginia Wilcox-Gok. 2007. Do the sick retire early? chronic illness, asset accumulation and early retirement. *Applied Economics* 39 (13-15): 1921–1936.

Minino, Arialdi M., Kenneth D. Kochanek, and Jiaquan Xu. 2008. Deaths: final data for 2008.

Murphy, Sherry L. 1998. Deaths: final data for 1998.

Murphy, Sherry L., Kenneth D. Kochanek, and Jiaquan Xu. 2010. Deaths: preliminary data for 2010.

Murphy, Sherry L., et al. 2000. Deaths: final data for 2000.

Murphy, Sherry L., et al. 2002. Deaths: final data for 2002.

Murphy, Sherry L., et al. 2004. Deaths: final data for 2004.

Murphy, Sherry L., et al. 2006. Deaths: final data for 2006.

National Center for Health Statistics. 1992. Vital statistics of the United States, 1992.

———. 1994. Vital statistics of the United States, 1994.

Nyce, Steven, et al. 2011. Does retiree health insurance encourage early retirement? *National Bureau of Economic Research Working Paper Series* No. 17703.

Peters, Kimberley D., Kenneth D. Kochanek, and Sherry L. Murphy. 1996. Deaths: final data for 1996.

Scandlen, Greg. 2001. Helping laid-off workers keep insurance. *Hartlander.* http://news.heartland.org/newspaper-article/2001/11/01/helping-laid-workers-keep-insurance.

Social Security Administration. 2011. *Disability benefits, ssa publication no. 05-10029.*

St.Clair, Patricia, et al. 2011. *RAND HRS data documentation, version L.*

U.S. Census Bureau. 2012. *Health insurance historical tables.* http://www.census.gov/hhes/www/hlthins/data/historical/index.html.

APPENDIX

A Dependent Variables

Table 1 displays the means, standard errors and percentile values for the two dependent variables. Figures 1 and 2 display the distribution of the differences between the highest and the lowest value of the dependent variable for each respondent.

Table 1: Means and distribution of dependent variables

Statistic	R	P(62)
Mean	65.55	49.39
Standard error	5.48	38.84
25th percentile	62.00	0.00
50th percentile	65.00	50.00
75th percentile	68.00	90.00

Figure 1: Histogram of planned retirement age differences

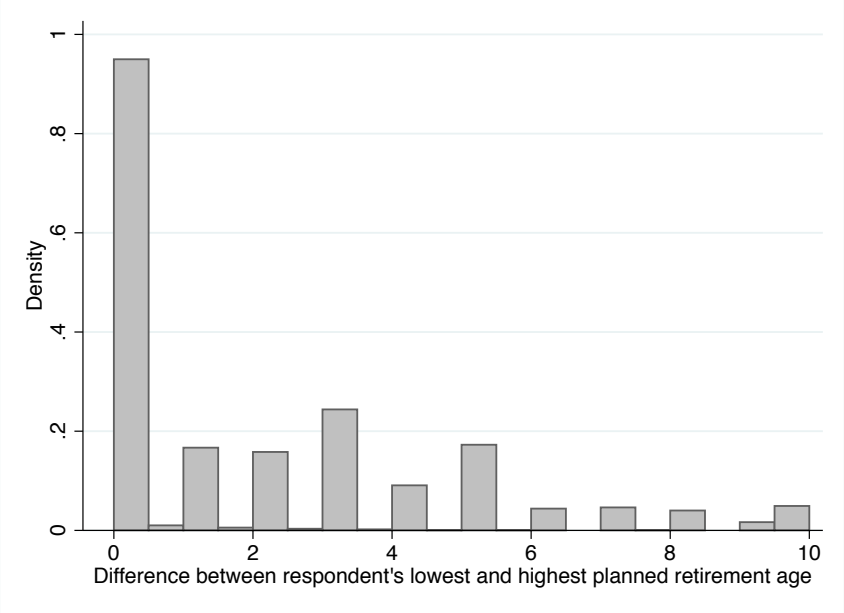
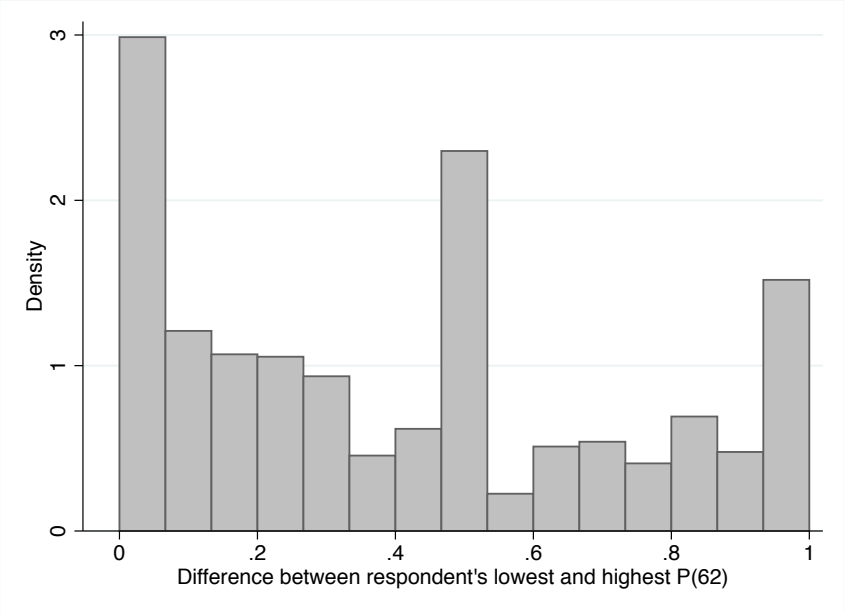


Figure 3 displays the differences between the planned and actual retirement age of individuals. It is important that the distribution is centered around zero. Table 2 shows the distribution

Figure 2: Histogram of the differences in probability of working full-time after the age of 62



in a more detailed way. The mean actual retirement age among those who stated a low probability (20% and below) of working full time after of 62 was at 61.39 years, while it was at 62.43 and 63.42 for those with probabilities of full-time work between 20% and 80% and those with probabilities of 80% and above respectively.

In an attempt to partially explain the discrepancy between planned and actual retirement age of individuals, a simple linear regression was run with the difference between the planned and actual retirement age as the dependent variable. The results, which can be seen in table 3, suggest that both health, health insurance and employment characteristics play a role in individuals' retirement predictions and the failure to predict retirement accurately. When it comes to health, worse health leads to an actual retirement age being much lower than that which was planned[39]. Having a retiree health insurance, a full time job, higher income and pension income seems to contribute to an actual retirement age being higher than what an individual had planned.

[39]The comparison group in the case of health status is Excellent health status. Activities of daily living score (ADL) is compared to those who had no difficulties with any of the daily tasks used to construct this score.

Figure 3: Histogram of the differences between planned and actual retirement age

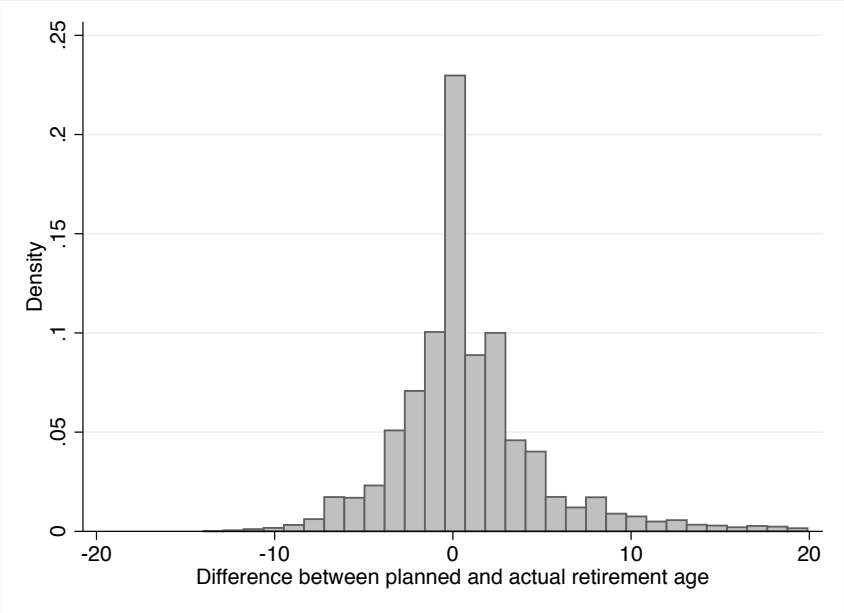


Table 2: Differences between planned and actual retirement age

Planned − actual retirement age	Percentage of sample (in %)
$= 0$ (Exact prediction)	4.63
> 0 (Earlier retirement)	46.48
< 0 (Later retirement)	48.89
Within 0.5 years	26.00
Between 0.51 and 1 year	11.50
Between 1.01 and 2 years	14.82
Between 2.01 and 3 years	16.38
More than 3 years	31.30

Table 3: Analysis of the main determinants of the differences between the planned and actual retirement age

Difference between the planned and actual retirement age	β	Standard error
Health		
Very good	0.311**	(0.150)
Good	0.347**	(0.158)
Fair	0.965***	(0.251)
Poor	1.377**	(0.623)
ADL 1	0.358	(0.428)
ADL 2	0.930	(0.904)
ADL 3	4.440*	(2.324)
ADL 4	5.639***	(2.087)
ADL 5	17.468**	(7.817)
Negative feelings score	0.026	(0.044)
Probability of living past age of 75	0.171	(0.219)
Health insurance		
Retiree HI available (any type)	−0.326***	(0.116)
Respondent now: government	−0.005	(0.246)
Respondent now: dependent	−0.022	(0.120)
Job characteristics		
Works full-time	−0.446*	(0.240)
Log of annual income	−0.553***	(0.097)
Average weekly hours of work	0.018**	(0.007)
Has pension	−1.192***	(0.160)
Observations	6044	
R^2	0.170	

Heteroskedasticity-robust standard errors in parentheses
** p <0.10, ** p <0.51, *** p <0.01*
Selected coefficients are reported. Other variables in the regression included household wealth, having government health insurance, demographic characteristics such as age, gender, marital status, race and religion. ADL stands for the score for activities of daily living (see appendix B for a definition). High score indicates high difficulties in performing simple daily tasks, low score indicates low difficulties (no difficulties omitted).

B Main Explanatory Variables: Health

Table 4: Description of health variables

Variable name	Description
Self-reported health status	
Self-reported health status	Scored as Excellent (1), Very good (2), Good (3), Fair (4), and Poor (5).
Activities of daily living score	This score is based on an individual's abilities to bathe, dress, eat, get in/out of bed, and walk across a room. Higher score means bigger problems with performing these activities. Some activities have been also included individually, but are not included in regressions with activities of daily living score.
Walking across a room	Whether an individual has any difficulty walking across a room.
Walking a block	Whether an individual has any difficulty walking a block.
Walking several blocks	Whether an individual has any difficulty walking several blocks.
Getting in and out of bed	Whether an individual has any difficulty getting in/out of bed.
Walking up a flight of stairs	Whether an individual has any difficulty walking up a flight of stairs.
Walking up several flights of stairs	Whether an individual has any difficulty walking up several flights of stairs.
Lifting 10 pounds	Whether an individual has any difficulty lifting 10 lbs. of weight.
Picking a small object (dime)	Whether an individual has any difficulty picking up a dime.
Negative feelings score	A summary measure of an individual's feelings during the week prior to the interview. Higher score means a higher level of negative feelings.
Probability of living past age of 75	The self reported probability of living past the age of 75.
Health limits work	Whether current health condition limits the amount/type of work one can do.
Medical care use	
Nights in hospital	The number of nights that an individual has spent in the hospital since the last interview.
Nights in nursing home	The number of nights that an individual has spent ina nursing home since the last interview.
Doctor visits	Number of times an individual visited a doctor since the last interview.
Prescription drugs	Whether regularly taking prescription drugs.
Outpatient surgery	Whether an individual had an outpatient surgery since the last interview.
Dentist services	Whether an individual used the dentist's services since the last interview.
Special facilities	Whether an individual used any special medical facilities since the last interview.
Illness occurrence	
Has high blood pressure	Whether an individual had high blood pressure at time of interview.
Has diabetes	Whether an individual had diabetes at time of interview.
Has cancer	Whether an individual has had cancer at any point since last interview or at time of interview.
Has lung disease	Whether an individual had a chronic lung disease (except asthma) at time of interview.
Has heart problems	Whether an individual had a heart disease (heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems) at time of interview or since last interview.
Has had a stroke	Whether an individual has had a stroke or transient ischemic attack since last interview.
Has psychological problems	Whether an individual had emotional, nervous, or psychiatric problems since last interview.
Has arthritis	Whether an individual had arthritis or rheumatism at time of interview or since last interview.
Medical expenditures	
Total medical OOP exp. (in $1,000)	Total out of pocket medical expenditures for the past two years (past one year in wave 1 interview) in thousands of dollars.
Total medical exp. (in $1,000)	Total estimated medical expenditures for the past two years (past 12 months in wave 1 interview) by brackets in thousands of dollars. This measure includes health care procedures paid for by health insurance and is available for waves 1 - 6 only.

Table 5: Descriptive statistics of health measures by self-reported health status

	Excellent	Very good	Good	Fair	Poor
Self-reported health					
Activities of daily living score	0.009	0.017	0.056	0.200	0.576
	(0.109)	(0.154)	(0.291)	(0.591)	(1.009)
Walking across a room	0.001	0.003	0.009	0.029	0.101
	(0.033)	(0.052)	(0.092)	(0.168)	(0.301)
Walking a block	0.005	0.011	0.034	0.098	0.270
	(0.071)	(0.104)	(0.182)	(0.298)	(0.444)
Walking several blocks	0.026	0.058	0.142	0.308	0.575
	(0.159)	(0.235)	(0.349)	(0.462)	(0.495)
Getting in and out of bed	0.005	0.007	0.022	0.072	0.180
	(0.068)	(0.086)	(0.147)	(0.259)	(0.384)
Sitting for two hours	0.090	0.115	0.170	0.278	0.415
	(0.286)	(0.319)	(0.376)	(0.448)	(0.493)
Getting up from a chair	0.115	0.200	0.313	0.462	0.622
	(0.318)	(0.400)	(0.464)	(0.499)	(0.485)
Walking up a flight of stairs	0.014	0.027	0.077	0.181	0.401
	(0.115)	(0.163)	(0.267)	(0.385)	(0.490)
Walking up several flights of stairs	0.128	0.232	0.387	0.588	0.756
	(0.334)	(0.422)	(0.487)	(0.492)	(0.430)
Lifting 10 pounds	0.037	0.060	0.119	0.241	0.468
	(0.188)	(0.237)	(0.324)	(0.428)	(0.499)
Picking up a small object (dime)	0.007	0.015	0.028	0.056	0.119
	(0.082)	(0.123)	(0.164)	(0.231)	(0.324)
Negative feelings score	0.608	0.846	1.336	2.297	3.582
	(1.178)	(1.424)	(1.797)	(2.283)	(2.475)
Probability of living past age of 75	0.773	0.706	0.630	0.533	0.410
	(0.227)	(0.240)	(0.276)	(0.315)	(0.340)
Health limits work	0.023	0.045	0.107	0.236	0.492
	(0.150)	(0.208)	(0.310)	(0.425)	(0.500)
Medical care use					
Nights in hospital	0.231	0.385	0.840	1.521	2.990
	(1.597)	(2.596)	(6.217)	(5.223)	(8.547)
Nights in nursing home	0.002	0.048	0.142	0.029	0.100
	(0.141)	(4.449)	(13.454)	(0.991)	(2.061)
Doctor visits	3.674	5.242	7.013	10.466	15.850
	(6.063)	(7.811)	(11.609)	(18.081)	(22.624)
Prescription drugs	0.416	0.582	0.676	0.767	0.829
	(0.493)	(0.493)	(0.468)	(0.423)	(0.377)
Outpatient surgery	0.142	0.161	0.174	0.191	0.226
	(0.349)	(0.367)	(0.379)	(0.393)	(0.419)
Dentist services	0.811	0.776	0.673	0.579	0.515
	(0.392)	(0.417)	(0.469)	(0.494)	(0.500)

Continued on next page

Continued from previous page

	Excellent	Very good	Good	Fair	Poor
Special facilities	0.028	0.042	0.051	0.064	0.083
	(0.165)	(0.200)	(0.221)	(0.245)	(0.277)
Illness occurrence					
Has high blood pressure	0.176	0.337	0.465	0.562	0.614
	(0.380)	(0.473)	(0.499)	(0.496)	(0.487)
Has diabetes	0.023	0.057	0.138	0.241	0.315
	(0.149)	(0.231)	(0.345)	(0.427)	(0.465)
Has cancer	0.034	0.050	0.065	0.086	0.130
	(0.182)	(0.219)	(0.247)	(0.281)	(0.337)
Has lung disease	0.016	0.030	0.059	0.109	0.185
	(0.125)	(0.171)	(0.235)	(0.312)	(0.388)
Has heart problems	0.027	0.062	0.110	0.183	0.246
	(0.161)	(0.242)	(0.313)	(0.387)	(0.431)
Has had a stroke	0.005	0.011	0.018	0.039	0.070
	(0.072)	(0.103)	(0.131)	(0.193)	(0.255)
Has psychological problems	0.048	0.084	0.118	0.192	0.293
	(0.213)	(0.277)	(0.323)	(0.394)	(0.456)
Has arthritis	0.201	0.337	0.444	0.544	0.638
	(0.401)	(0.473)	(0.497)	(0.498)	(0.481)
Medical expenditures					
Total medical OOP exp. (in $1,000)	1.270	1.610	2.074	2.964	4.179
	(3.017)	(3.648)	(5.821)	(7.106)	(10.868)
Total medical exp. (in $1,000)	3.495	4.645	6.623	11.168	18.895
	(11.943)	(16.227)	(20.280)	(32.781)	(47.040)

Mean coefficients with standard deviations in parentheses. Number of observations may differ depending on the variable.

C Main Explanatory Variables: Health Insurance

Table 6 gives formal definitions to the health insurance types encountered in this study. Table 7 displays the process of sorting individuals into the second and third type of health insurance groups (division by the health insurance type in case of early retirement), one that I find the most meaningful when looking at how health impacts retirement decision. The advantage of this sorting method is that we avoid double-counting individuals and also look at what is relevant to the respondent when deciding when and whether to retire[40]. Furthermore, tables 8, 9 and 10 display the means of health insurance measures by health status as well as the means of different health indicators displayed by health insurance types.

[40]Even though many previous studies look at health insurance at the time of the survey or the availability of retiree health insurance, it seems clear that when individuals make retirement decisions, they consider whether they retain health insurance after their early retirement. Looking only at the availability of retiree health insurance is a highly simplistic approach, which may be inappropriate especially in case of couples.

Table 6: Description of health insurance variables

Health insurance type	Explanation
Division by the availability of retiree health insurance	
Retiree	The individual has health insurance and available retiree insurance through his/her employer.
Employer but no retiree	The individual has health insurance through his/her employer but the employer offers no retiree health insurance.
Government-provided	The individual is receiving coverage through either Medicaid or Medicare or other government programs.
Other	The individual reports having zero insurance plans, the individual is insured through an individual / alternative policy or the individual is insured as a dependent.
Division by the health insurance type in case of early retirement (no spouse)	
Single — insured	The individual is insured through a retiree insurance policy or is receiving coverage through a government program.
Single — uninsured	The individual is not insured through an employer's policy and thus depends on own assets for health care financing.
Division by the health insurance type in case of early retirement (married)	
Married — insured	In case of respondent's early retirement, both the respondent and spouse are insured.
Married — insured/uninsured	In case of respondent's early retirement, the respondent remains insured but spouse is uninsured*.
Married — uninsured/insured	In case of respondent's early retirement, the respondent is uninsured* while the spouse is insured.
Married — uninsured	In case of respondent's early retirement, both the respondent and spouse are left uninsured*.

*Uninsured means that the individual depends on own assets for health care financing.

Table 7: Early-retirement health insurance options of couples

Spouse: current health insurance type		Respondent: current source of health insurance					
		Employer			Spouse	Government	Self-insured
		With RH both	With RH self	No RH			
Employer	Dep. possible	1	1	1	1	1	N/A
	Dep. not possible	1	1	3	NA	1	3
Dependent	Employer av. - dep. possible	1	1	1	N/A	N/A	N/A
	Employer av. - dep. not possible	1	1	3	N/A	N/A	N/A
	Employer not available	1	2	4	N/A	N/A	N/A
Government	Medicaid	1	1	1	N/A	1	1
	Medicare (and other)	1	1	3	N/A	1	3
Self-insured		1	2	4	N/A	2	4

Post early retirement health insurance groups: (1) Both respondent and spouse are insured, (2) Respondent insured, spouse self-insured, (3) Respondent self-insured, spouse insured, (4) Both respondent and spouse are self-insured. Self-insured, for the purpose of this table, indicates a state in which an individual needs to provide for his or her own insurance, either through a private insurance company or pay for health care with own assets. N/A refers to a health insurance situation that is not possible. Additionally I make the following assumptions: (a) if a respondent is self-insured, a spouse's employer does not offer health insurance for dependents; (b) group 1, if a respondent reports being dependent on spouse for health insurance;

Table 8: Descriptive statistics of health insurance measures by self-reported health status

	Excellent	Very good	Good	Fair	Poor
Division by the availability of retiree health insurance					
Retiree	0.352	0.350	0.327	0.269	0.227
	(0.478)	(0.477)	(0.469)	(0.443)	(0.419)
Employer but no retiree	0.303	0.310	0.307	0.302	0.277
	(0.459)	(0.462)	(0.461)	(0.459)	(0.448)
Government-provided	0.026	0.023	0.030	0.046	0.068
	(0.160)	(0.149)	(0.170)	(0.210)	(0.252)
Other	0.320	0.318	0.337	0.384	0.429
	(0.466)	(0.466)	(0.473)	(0.486)	(0.495)
Division by the health insurance type in case of early retirement (no spouse)					
Single — insured	0.588	0.586	0.563	0.505	0.464
	(0.492)	(0.493)	(0.496)	(0.500)	(0.499)
Single — uninsured	0.412	0.414	0.437	0.495	0.536
	(0.492)	(0.493)	(0.496)	(0.500)	(0.499)
Division by the health insurance type in case of early retirement (married)					
Married — insured	0.513	0.505	0.476	0.409	0.367
	(0.500)	(0.500)	(0.499)	(0.492)	(0.482)
Married — insured/uninsured	0.074	0.081	0.088	0.096	0.097
	(0.262)	(0.272)	(0.283)	(0.295)	(0.296)
Married — uninsured/insured	0.191	0.205	0.191	0.186	0.155
	(0.393)	(0.404)	(0.393)	(0.389)	(0.362)
Married — uninsured	0.221	0.209	0.245	0.309	0.381
	(0.415)	(0.406)	(0.430)	(0.462)	(0.486)
Observations	50924				

Mean coefficients with standard deviations in parentheses.

Table 9: Means of health variables by current health insurance type

	Retiree	Employer, no retiree	Government	Other
Self-reported health status				
Self-reported health status	2.320	2.395	2.645	2.466
	(0.965)	(0.996)	(1.124)	(1.036)
Activities of daily living score	0.047	0.059	0.106	0.071
	(0.281)	(0.323)	(0.432)	(0.363)
Walking across a room	0.007	0.009	0.022	0.010
	(0.084)	(0.096)	(0.146)	(0.102)
Getting in and out of bed	0.019	0.021	0.035	0.027
	(0.135)	(0.143)	(0.185)	(0.163)
Sitting for two hours	0.155	0.137	0.209	0.159
	(0.362)	(0.344)	(0.407)	(0.366)
Getting up from a chair	0.248	0.256	0.315	0.261
	(0.432)	(0.436)	(0.465)	(0.439)
Lifting 10 pounds	0.097	0.098	0.163	0.109
	(0.296)	(0.297)	(0.370)	(0.311)
Picking up a small object (dime)	0.023	0.022	0.033	0.027
	(0.151)	(0.146)	(0.178)	(0.162)
Negative feelings score	1.014	1.146	1.561	1.306
	(1.589)	(1.730)	(2.058)	(1.883)
Medical care use				
Nights in hospital	0.656	0.715	1.504	0.607
	(3.290)	(3.805)	(16.432)	(3.440)
Nights in nursing home	0.006	0.021	0.164	0.159
	(0.392)	(1.139)	(6.268)	(13.521)
Doctor visits	6.058	6.988	7.869	5.776
	(9.931)	(11.543)	(12.709)	(11.845)
Prescription drugs	0.636	0.634	0.646	0.553
	(0.481)	(0.482)	(0.478)	(0.497)
Outpatient surgery	0.187	0.178	0.165	0.137
	(0.390)	(0.383)	(0.372)	(0.343)
Illness occurrence				
Has high blood pressure	0.388	0.377	0.453	0.354
	(0.487)	(0.485)	(0.498)	(0.478)
Has diabetes	0.101	0.103	0.148	0.097
	(0.302)	(0.304)	(0.355)	(0.296)
Has cancer	0.056	0.060	0.075	0.056
	(0.229)	(0.237)	(0.264)	(0.230)
Has lung disease	0.049	0.048	0.065	0.047
	(0.217)	(0.214)	(0.247)	(0.211)
Has heart problems	0.094	0.096	0.129	0.071
	(0.292)	(0.294)	(0.335)	(0.257)
Has had a stroke	0.016	0.016	0.027	0.016
	(0.125)	(0.126)	(0.163)	(0.124)
Has psychological problems	0.088	0.105	0.156	0.115
	(0.283)	(0.306)	(0.363)	(0.319)
Has arthritis	0.377	0.370	0.420	0.367
	(0.485)	(0.483)	(0.494)	(0.482)
Medical expenditures				
Total medical OOP exp. (in $1,000)	1.631	2.041	1.556	2.057
	(3.469)	(5.011)	(5.044)	(6.300)
Total medical exp. (in $1,000)	5.941	6.711	8.258	5.296
	(18.796)	(22.668)	(25.156)	(19.853)

Mean coefficients with standard deviations in parentheses. Data in column 1 represent those who are in the Single — insured category, data in column 4 represent those in the Single — uninsured category.

Table 10: Means of health variables by early-retirement health insurance type

	Married — insured	M. — insured/ uninsured	M. — uninsured/ insured	M. — uninsured
Self-reported health status				
Self-reported health status	2.332	2.484	2.376	2.533
	(0.976)	(1.016)	(0.981)	(1.067)
Activities of daily living score	0.050	0.069	0.059	0.080
	(0.295)	(0.341)	(0.323)	(0.389)
Walking across a room	0.007	0.012	0.011	0.011
	(0.086)	(0.108)	(0.102)	(0.106)
Getting in and out of bed	0.019	0.024	0.023	0.029
	(0.137)	(0.152)	(0.149)	(0.168)
Sitting for two hours	0.150	0.149	0.151	0.159
	(0.357)	(0.356)	(0.358)	(0.366)
Getting up from a chair	0.246	0.267	0.267	0.268
	(0.430)	(0.442)	(0.442)	(0.443)
Lifting 10 pounds	0.090	0.125	0.104	0.122
	(0.286)	(0.330)	(0.306)	(0.327)
Picking up a small object (dime)	0.022	0.024	0.024	0.029
	(0.147)	(0.154)	(0.154)	(0.169)
Negative feelings score	0.980	1.374	1.080	1.533
	(1.584)	(1.846)	(1.672)	(2.025)
Probability of living past age of 75	0.673	0.684	0.676	0.664
	(0.266)	(0.277)	(0.268)	(0.287)
Medical care use				
Nights in hospital	0.703	0.859	0.610	0.635
	(5.199)	(4.292)	(2.959)	(3.800)
Nights in nursing home	0.011	0.075	0.062	0.178
	(0.896)	(3.752)	(5.432)	(15.256)
Doctor visits	6.234	7.168	6.796	5.725
	(10.583)	(10.589)	(12.426)	(11.497)
Prescription drugs	0.628	0.628	0.624	0.543
	(0.483)	(0.483)	(0.484)	(0.498)
Outpatient surgery	0.183	0.173	0.167	0.131
	(0.387)	(0.378)	(0.373)	(0.337)
Illness occurrence				
Has high blood pressure	0.374	0.435	0.352	0.376
	(0.484)	(0.496)	(0.478)	(0.484)
Has diabetes	0.100	0.118	0.097	0.103
	(0.300)	(0.322)	(0.296)	(0.304)
Has cancer	0.054	0.066	0.064	0.057
	(0.226)	(0.249)	(0.244)	(0.232)
Has lung disease	0.046	0.060	0.045	0.052
	(0.210)	(0.238)	(0.208)	(0.222)
Has heart problems	0.099	0.079	0.079	0.075
	(0.299)	(0.270)	(0.270)	(0.263)
Has had a stroke	0.017	0.017	0.014	0.016
	(0.130)	(0.131)	(0.118)	(0.125)
Has psychological problems	0.088	0.128	0.107	0.126
	(0.283)	(0.334)	(0.309)	(0.332)
Has arthritis	0.360	0.415	0.386	0.373
	(0.480)	(0.493)	(0.487)	(0.484)
Medical expenditures				
Total medical OOP exp. (in $1,000)	1.747	1.749	2.106	2.081
	(3.858)	(3.961)	(4.817)	(7.290)
Total medical exp. (in $1,000)	6.193	6.951	6.285	5.073
	(19.588)	(24.443)	(22.453)	(19.194)

Mean coefficients with standard deviations in parentheses.

D Control Variables

This appendix carefully lists the many control variables used in the various regressions of this study. Furthermore, tables 12 and 13 give the means of these control variables for sample health insurance types.

Table 11: Comprehensive listing of control variables

Category	Control variables
Respondent	
Demographic	Age, gender, race and ethnicity, education, religion, probability of living past the age of 75 (from Vital Statistics), marital status, number of children, income risk aversion, whether one has good friends in the neighborhood, HRS cohort.
Parental information	Whether mother alive, whether father alive.
Employment	Employment status, veteran status, annual income, average weekly hours, experience, employer size, occupation and industry groups, whether covered by a union contract, physical effort and stress levels, job satisfaction, adequacy of salary report, whether ever reported staying in current job because of health insurance, weeks of paid vacation.
Marriage	Whether lives in a couple household, length of current marriage, how enjoyable is the time spent together with spouse, whether free time is spent with spouse.
Retirement funds	Net household wealth, whether has pension, whether was previously offered a retirement incentive.
Spouse (if applicable)	
Demographic	Age, gender, race and ethnicity, education, religion, probability of living past the age of 75 (from Vital Statistics), whether receiving SSDI, income risk aversion.
Parental information	Whether mother alive, whether father alive.
Employment	Employment status, veteran status, annual income, experience, whether has pension income, whether has pension, planned retirement age, actual age of retirement.
Region level	
	Census region or division, unemployment rate in the Census Division, region level retirement and health insurance policies, gross domestic product.

Not all of these variables will be used in all regressions as some have a high number of missing observations.

Table 12: Means of main control variables by health insurance type

	Retiree	No retiree	Married - insured	Married - uninsured
Demographic				
Age at interview	56.550	55.988	56.094	56.363
	(4.821)	(5.271)	(5.025)	(5.126)
Female	0.491	0.576	0.443	0.606
	(0.500)	(0.494)	(0.497)	(0.489)
Individual is white/ Caucasian	0.791	0.804	0.830	0.752
	(0.407)	(0.397)	(0.376)	(0.432)
Individual is black/African American	0.170	0.132	0.123	0.173
	(0.376)	(0.339)	(0.329)	(0.378)
Individial is Hispanic	0.017	0.033	0.023	0.044
	(0.130)	(0.180)	(0.149)	(0.206)
Individual is of other race/ ethnicity	0.022	0.030	0.024	0.031
	(0.146)	(0.170)	(0.153)	(0.172)
Years of education	13.345	12.783	13.311	12.217
	(2.617)	(3.077)	(2.749)	(3.392)
Individual is protestant	0.654	0.601	0.625	0.596
	(0.476)	(0.490)	(0.484)	(0.491)
Individual is catholic	0.250	0.283	0.270	0.284
	(0.433)	(0.450)	(0.444)	(0.451)
Individual is Jewish	0.016	0.022	0.020	0.018
	(0.124)	(0.147)	(0.140)	(0.134)
Probability of living past age of 75	0.681	0.668	0.673	0.664
	(0.266)	(0.276)	(0.266)	(0.287)
Marital status	0.791	0.791	0.978	0.452
	(0.406)	(0.407)	(0.147)	(0.498)
Total number of children	3.026	3.157	3.164	3.040
	(1.957)	(2.009)	(1.916)	(2.117)
Income risk aversion	3.334	3.227	3.310	3.210
	(1.198)	(1.453)	(1.257)	(1.505)
Individual has good friends in the neighborhood	0.612	0.603	0.605	0.606
	(0.487)	(0.489)	(0.489)	(0.489)
Parental information				
Mother is alive	0.446	0.445	0.457	0.425
	(0.497)	(0.497)	(0.498)	(0.494)
Father is alive	0.197	0.213	0.217	0.190
	(0.398)	(0.410)	(0.412)	(0.392)
Employment				
Works full-time	0.906	0.764	0.894	0.747
	(0.291)	(0.425)	(0.308)	(0.435)
Works part-time	0.081	0.193	0.090	0.197
	(0.273)	(0.395)	(0.286)	(0.398)
Unemployed	0.012	0.043	0.016	0.057
	(0.111)	(0.203)	(0.125)	(0.231)
Veteran status	0.259	0.182	0.273	0.143
	(0.438)	(0.386)	(0.446)	(0.350)
Observations	16869	34064	24636	12122

Mean coefficients, standard deviations in parentheses

Table 13: Means of main control variables by health insurance type - continued

	Retiree	No retiree	Married - insured	Married - uninsured
Employment				
Log of annual income	3.477	3.143	3.500	2.939
	(0.821)	(1.044)	(0.856)	(1.061)
Average weekly hours of work	42.420	40.761	42.789	41.043
	(10.895)	(14.687)	(11.676)	(15.378)
Experience	34.258	31.657	33.821	31.077
	(8.982)	(10.834)	(9.320)	(11.370)
Employer size	458.376	262.591	454.319	186.031
	(1725.721)	(1373.795)	(1767.925)	(963.321)
Occupation	5.830	6.038	5.892	6.519
	(4.932)	(4.683)	(4.974)	(4.678)
Industry	8.012	7.904	7.764	7.733
	(3.983)	(3.749)	(3.946)	(3.740)
Union contract	0.286	0.139	0.248	0.119
	(0.452)	(0.346)	(0.432)	(0.324)
Job requires a lot of physical effort*	2.158	2.295	2.153	2.438
	(1.098)	(1.137)	(1.099)	(1.156)
Job involves a lot of stress*	2.849	2.778	2.853	2.749
	(0.804)	(0.821)	(0.798)	(0.832)
Individual enjoys work*	2.988	2.892	2.907	2.926
	(1.175)	(1.216)	(1.158)	(1.218)
Individual finds salary adequate*	2.843	2.729	2.844	2.637
	(0.726)	(0.809)	(0.736)	(0.825)
Reported staying in current job because of health insurance	0.340	0.312	0.357	0.304
	(0.474)	(0.463)	(0.479)	(0.460)
Weeks of paid vacation	3.937	2.917	3.757	2.583
	(4.653)	(4.638)	(4.460)	(4.886)
Marriage				
Couple household	0.792	0.791	0.978	0.453
	(0.406)	(0.407)	(0.145)	(0.498)
Length of current marriage	27.532	26.759	27.236	26.010
	(11.362)	(11.629)	(11.345)	(12.037)
Time spent with spouse is enjoyable*	3.149	3.102	3.145	3.029
	(0.698)	(0.725)	(0.702)	(0.749)
Free time is spent together with spouse*	2.388	2.404	2.397	2.409
	(0.734)	(0.734)	(0.731)	(0.763)
Retirement funds				
Net household wealth (in $1,000)	307.231	358.850	385.601	274.098
	(1122.047)	(1304.414)	(1355.289)	(1271.633)
Has pension	0.800	0.476	0.759	0.351
	(0.400)	(0.499)	(0.428)	(0.477)
Previously offered retirement incentive	0.066	0.018	0.051	0.013
	(0.249)	(0.132)	(0.220)	(0.112)
Observations	16869	34064	24636	12122

Mean coefficients, standard deviations in parentheses
* Rated on a scale from 1 to 4, 1 being a strong disagreement (statement is false), 4 being a strong agreement (statement is true).

E Details of the First Step Estimation

As described in equation 6, objective health and demographic variables were regressed on the subjective self-reported health status using simple ordered probit model. After the coefficients were estimated, the probability p_i of each possible health status i was calculated. Following this step, I then estimated two alternative forms of fitted values: (a) a health status selected based on the highest probability; and (b) a probability of poor or fair health calculated as the sum of the two respective probabilities.