J Labor Res (2018) 39:56–98
DOI 10.1007/s12122-017-9255-6

Effect of Health on Retirement of Older Americans: a Competing Risks Study

Subhasree Basu Roy [1]

Published online: 1 August 2017

Abstract Retirement is an important event in the life of an individual. The decision to retire or exit from full-time employment may be motivated by several factors, including health. This paper explores the effect of both subjective and relatively more objective physical and mental health conditions on the probability of exit from full-time employment. Using longitudinal data on older Americans from ten waves of the Health and Retirement Study (1992–2010), eight health indices are created from a wide range of health measures by principal component analysis. The effect of these health indices on the time until exit from full-time employment is empirically examined in a proportional hazard model. Single and competing risk specifications are estimated that allow for multiple spells of full-time employment and control for unobserved heterogeneity. The main results suggest that better self-reported health decreases the likelihood of exit from full- time employment, while poor physical health (functional limitations factor) increases the likelihood of exit from full-time employment via complete retirement and disability. For mental health, I find that depression increases the likelihood of exit via complete retirement, part-time work and unemployment while cognitive disorders lead to an increase in likelihood of exit via the disability exit route. Hence, physical and mental health problems are both impediments to continued work. These results have implications for public policies targeted towards retaining older workers within the labor market.

Keywords Health indices · Principal component analysis · Retirement · Competing risks · Proportional hazard · Unobserved heterogeneity

JEL Classification I10 · J14 · J26 · C41

✉ Subhasree Basu Roy
SBasuRoy@missouristate.edu

[1] Department of Economics, Missouri State University, 901 S. National Ave, Springfield, MO 65897, USA

Introduction

Retirement is an important event in the life of an individual. Among many other factors (socio-economic and financial), individual health status may have a strong influence on retirement decisions. For the last few decades, population aging has foreshadowed serious policy problems throughout the world, including within the United States. With the rapid rise in aging population in U.S, an increasingly high proportion of individuals are fast approaching their full retirement age (65 years). According to Social Security Administration (Office of Policy), the proportion of people over 65 years of age in 2009 was 12.9% as compared to 8% in 1950. Over next few decades, as the baby boom generation (born during 1946–1964) enters their elderly years the proportion of individuals over 65 years of age is projected to rise to 20% (in 2040).

These demographic changes suggest financing challenges for transfer programs such as the Social Security program. As a result, policymakers may promote policies designed to retain productive older workers in the workforce in order to defer their Social Security payments. The success of such policies depends in part on the ability to identify the key determinants of a worker's decision to retire and the magnitude of their impact. One obvious factor that plays into a worker's decision to continue working or retire is their health status. Policies that improve the health of workers may encourage them to continue working and defer the start of their Social Security payments.[1] The purpose of this paper is to examine the role of physical and mental health conditions in determining the duration of full-time employment for older Americans. In other words, this paper measures the extent to which health influences the decision to retire among older U.S. workers.

There is a literature focused on identifying the causal effect of health on retirement (Anderson and Burkhauser 1985; Bazzoli 1985; Bound 1991; Bound et al. 1998; Disney et al. 2006; Dwyer and Mitchell 1999; Lindeboom and Kerkhofs 2002). These studies primarily used self-reported subjective measures of health. Such measures of health may be plagued with problems that lead to bias. First, self-reported measures of health are based on subjective judgments and there is no reason to believe that these judgments are comparable across individuals. Second, since poor health may represent a legitimate reason for a person of working age to be outside the labor force, respondents who are not working may cite health problems as a way to rationalize behavior (the "justification hypothesis"). A final issue is that many papers in the literature are forced to rely on relatively short panels due to limitations in data availability at the time of the studies.

This paper adds to the existing literature by using duration analysis on a panel dataset of older Americans (those at least 50 years old and working full-time in 1992) and utilizes a wider variety of health indices to estimate the impact of health on the duration of full-time employment. Unlike much of the previous literature, the long panel nature of the Health and Retirement Study dataset is exploited here, making it possible to observe more cases of actual retirement than retirement plans

[1] However, such investments might come at an "unanticipated cost" of extending the individual's lifespan, so that the government may end up paying more in social security over all. Boskin et al. (1987); Coile et al. (2002) and Gruber and Wise (2005) explore this area, which is beyond the focus of this paper.

(expectations) and the potential to observe multiple spells of employment over a 20-year timeframe.[2] Excluding younger individuals and those initially working less than full-time reduces concern about the justification bias. The issue of subjectivity in health outcomes is addressed by constructing eight relatively objective health indices (factors) through principal component analysis (PCA) that are based on a broad range of subjective as well as objective health measures.[3] Unlike previous studies, it is possible to observe how different health measures load in these indices. Besides the health indices created by PCA, changes in physical and mental health between consecutive waves are considered.

The main results for the overall study sample suggest that among physical health factors, an increase in the functional limitations factor increases likelihood of exit from full-time work by 18.88% overall, which is largely driven by exits via complete retirement or disability routes. However, the probability of exit for complete retirement is much larger in magnitude. On the other hand, cancer leads to a 7.71% decrease in likelihood of exit via complete retirement. Among mental health factors, an increase in cognitive disorders has no significant effect on exit via the complete retirement route but leads to a 1.14% increase in likelihood of exit via disability. An increase in depression factor leads to 9.06%, 3.04 and 0.90% increase in likelihood of exit via complete retirement, part-time work and unemployment routes respectively. An increase in the risky lifestyle behavior factor (smoking, drinking) leads to an increase in likelihood of exit via complete and partial retirement routes.

The rest of this paper is organized as follows: section 2 reviews previous literature and section 3 presents the data. A discussion of the empirical methodology is provided in section 4 followed by a discussion of results in section 5 and ending with a conclusion in section 6.

Literature Review

An early paper on this topic, Boskin (1977), does not find any significant effect of health on retirement, but a large effect of Social Security income. Unlike Boskin (1977), Quinn (1981) finds that the presence of a health limitation reduced the probability of labor force participation by 20 percentage points. Bazzoli (1985) looks at the determinants of early retirement and the impact of various measures of health. She also addresses the issue of the relative importance of health and economic factors in influencing early retirement. She finds economic factors rather than health factors play the major role in retirement decisions.

In attempting to identify the causal effect of health on retirement decisions, the use of subjective measures of health has been a focus of much attention in the literature (Anderson and Burkhauser 1985; Bazzoli 1985; Bound 1991; Bound et al. 1998; Disney et al. 2006; Dwyer and Mitchell 1999; Lindeboom and Kerkhofs 2002). Such measures of health may be plagued with problems that lead to bias. First, self-reported measures of health are based on subjective judgments (leading to bias) and there is no reason to believe that these judgments are comparable across individuals. The size of

[2] Since each wave is two years apart.

[3] Twenty eight different health measures have been used in PCA.

the bias present in self-reported health measures is documented in Benítez-Silva et al. (2004). Second, self-reported health may not be independent of labor market status. Third, since poor health may represent a legitimate reason for a person of working age to be outside the labor force, respondents who are not working may cite health problems as a way to rationalize their behavior ("justification hypothesis"). Fourth, for individuals for whom the financial rewards of continuing in the labor force are low there exists a financial incentive to report poor health as means of obtaining disability benefits. This is often cited as the "disability route" into retirement (Marmot et al. 2002; Riphahn 1999). For example, in a study of social security benefit programs in the Netherlands, Kerkhofs and Lindeboom (1995) show that recipients of disability insurance systematically overstate their health problems.

A large literature attempts to address this concern about subjective reports of health status. To mitigate response bias, authors have attempted to use arguably more objective measures of health, such as the observed future death of respondents (Anderson and Burkhauser 1985; Parsons 1982). Parsons (1982) and Stern (1989) find those who withdraw from the labor market are likely to cite poor health as the cause even if they are not in poor health, simply because they may be rewarded for doing so through eligibility for transfers. Their findings suggest that the traditional measure of self-reported health is endogenous due to "justification hypothesis." In other words, people who intend to keep working will downplay their health problems while ones who dislike work and wish to exit from the labor force will exaggerate their health problems. Dwyer and Mitchell (1999) implement an instrumental variable approach to deal with endogeneity using the first four waves of HRS. The authors used parental health and mortality, respondent's age, number of children, and BMI as instruments for self-assessed health. They do not find evidence to support the justification hypothesis. This could be because in the first four waves the HRS sample is still relatively young and does not yet consist of a majority of retired individuals.

Bound et al. (1998) examined the relationship between health and alternative labor force transitions like retirement, job change, and application for disability insurance. Their analysis not only considers health status, but also declines in health and its effect on the work behavior of individuals. According to the authors, retirement is often a last resort. Prior to such an outcome, workers may resort to increased effort, putting in more time, requesting a reduction in performance standards, or changing jobs in order to accommodate their physical limitations.

Using the first few waves of the HRS, McGarry (2004) models the labor market behavior and retirement probabilities of older workers prior to their eligibility for early retirement benefits and Social Security. She finds that changes in retirement expectations are driven more by health than economic variables. The effect of subjective measures of health is strong even when objective measures are included. Miah and Wilcox-Gök (2007) also use HRS data to determine how chronic illness affects asset accumulation and retirement. They find that the vast majority of the chronically ill population does not report their general health to be poor nor do they report functional limitations in activities of daily living. Nevertheless, their results indicate that chronic illness leads these people to accumulate fewer assets during their working years and consequently retire later.

While most studies in the existing literature use a fixed effect approach on panel data from different countries to investigate the effect of health on retirement, Meghir and Whitehouse (1997), Christensen and Kallestrup-Lamb (2012) and Siddiqui (1997) use duration models to study effect of poor health on labor force transitions for UK, British and Danish panel surveys respectively. In sum, these findings strongly suggest that poor health is a determinant of retirement or exit from the work force.

Data

The analysis presented in this paper exploits a long panel of data for Americans (1992–2010) from the Health and Retirement Study (HRS) conducted by the Institute for Social Research at the University of Michigan.[4] The HRS is an ongoing longitudinal survey, which began in 1992, and is conducted in biennial waves. Prior to 1998, the main HRS cohort included individuals born between 1931 and 1941, and another distinct cohort, the Study of Assets and Health Dynamics among the Oldest Old (AHEAD), included individuals born before 1924. Since 1998, the data for these two cohorts is collected jointly, and the sample frame has been expanded to include cohorts born between 1924 and 1930 and those born between 1942 and 1947. The HRS is administered for the specific purpose of studying life-cycle changes in health and economic resources, and includes detailed information on various subjective and objective health outcomes. In this paper, I focus on a sample of older individuals who were at least 50 years of age who were working full-time in 1992. The sample consists of 4128 individuals having multiple records that generate 15,442 person-wave observations. This was obtained by strictly dropping all cases with inconsistent or missing information for health measures and socio-demographic or economic variables.

Employment Spells

I consider two types of employment spells in my analysis. I start by focusing on initial employment spells by following the 4128 individuals in my sample starting in wave 1 (1992) over the subsequent waves until their first observed exit from full-time employment occurs (initial exit model).[5]At the end of each two year wave of the HRS, an individual who was working full-time in the previous wave could either continue working (and thus be treated as a right censored observation for that wave) or exit full-time work via one of the five different routes: complete retirement, partial retirement, part-time work, unemployment and disability. The exit routes are defined using the labor force participation variable.[6] Individuals leave full-

[4] The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. More information is available at: http://hrsonline.isr.umich.edu/

[5] Since I am focusing here on initial employment spells, any subsequent transitions back to full-time work are ignored. This implies I am considering retirement to be an absorbing state.

[6] In the HRS, individuals who report working 35 or more hours per week are considered full-time and those working less are considered part-time. This includes the hours and weeks worked in both the main and second job. The key HRS variable of interest here is the labor force participation variable (LFPV). If the respondent reports working full- time then their LFPV is set to that status. If he/she is working part-time and also reports retirement LFPV is set to partly retired. If there is no such reporting of retirement, then the variable is set to working part-time. If the respondent is neither working nor looking for work but there is reporting of retirement, then his LFPV is set to retired (completely retired). If retirement is not mentioned and a disabled employment status is given, then it is set to disabled. Otherwise, it is set to "not in the labor force." If the respondent is not working but is looking for a full-time work, labor force participation is set to unemployed. If he/she is looking for a part-time job and mentions retirement, then it is set to partly retired. Finally, individuals looking for a part-time job and not reporting retirement, have LFPV set to unemployed.

time employment through one of the five exit routes mentioned above or leave the survey via attrition or for some other reason, such as death. Those that remain in full-time employment or that exit the survey are treated as right censored employment spells.[7] Overall, 3.4% of the sample remains continuously full-time employed across all of the waves we observe. Since I am focusing here on initial employment spells, any subsequent transitions back to full-time work are ignored. This implies I am considering retirement to be an absorbing state.

The other type of employment spell I consider in my analysis is a subsequent employment spell (multiple exit model). In other words I drop the assumption that retirement is an absorbing state and allow members of the sample that retire to contribute additional spells of full-time employment if they re-enter the labor market. This adds 740 additional employment spells to the 4128 initial employment spells described above, for a total of 4868 full-time employment spells generated by my sample.[8]

Figure 1 depicts the baseline hazard rate for individuals to exit over time via any route in general and also via the five different exit routes. It is observed that the hazard rate for exiting full-time employment in general (via any route) cumulatively rises over time with a distinct peak occurring in Wave 7 (2004). This is largely driven by the rise in hazard rate of exiting via the complete retirement route which also peaks in Wave 7. This is probably because the individuals who are 50 years of age or older in 1992 become eligible for retirement benefits around the same time. The figure also indicates that the baseline hazard rate for the different exit routes is non-linear and not constant over time which calls for its parametric estimation using a suitable distribution.[9] The Kaplan-Meir survival estimates indicate that the probability of surviving in full-time employment declines over time. However, this decline is larger in the initial waves. This is true for all routes combined and the complete retirement and partial retirement routes. For the part-time work, unemployment and disability or not in labor force routes, the decrease in survival rate almost flattens out over time.

Health Measures

I use a wide range of health measures in this study in an attempt to address the concern that many health measures, such as self-reported health, are based on individual perceptions and may be plagued by misreporting and measurement error. Some of these measures are relatively more objective than others are and have not been used in

[7] The available HRS data allows me to follow the individuals through nine transitions: 1992–94, 1994–96, 1996–98, 1998–00, 2000–02, 2002–04, 2004–06, 2006–08, and 2008–10. The 4128 individuals who were at least 50 years old and worked full-time in 1992 are followed over the next 18 years (1992–2010). Between 1992 and 1994, 20 individuals leave the sample due to attrition (death or other reason), so 4108 individuals are "at risk" of retirement (exit) during 1994–1996. Among them 349 individuals already exit via complete retirement route by 1994. Another 148, 209, 70 and 67 individuals exit via partial retirement, part-time work, unemployment and disability/not in labor force routes respectively. Of the 3265 individuals who remain full-time employed in the sample in 1994, 27 are lost due to attrition, so only 3238 remain at the risk of retirement. Among them 372 individuals already exit via complete retirement route by 1996. Another 173, 20, 87 and 76 individuals exit via partial retirement, part-time work, unemployment and disability/not in labor force routes respectively. Similar pattern is observed between each two year time period. Finally, among the 4128 individuals only 142 remain full-time employed through all 18 years while the rest exited via one of the five routes or attrite.

[8] This is reported in Appendix, Table 8.

[9] The associated baseline hazard rates are reported in Appendix Table 9.

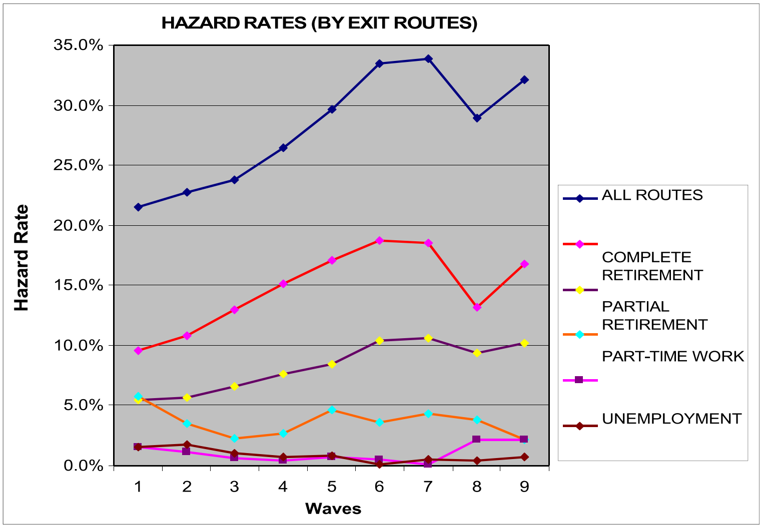

Fig. 1 Cumulative hazard rates for different exits (1992–2010)

the previous studies. Below I discuss these health measures by grouping them into four broad categories.

Self-reported health: This is the most subjective measure of health and has been widely used in existing studies. In the HRS, individuals may report their health as excellent, very good, good, fair and poor.

Physical health: Some of my measures of physical health have been used in prior studies, including counts of difficulties with activities of daily living (ADL) and diagnosed chronic conditions. The ADL difficulties include difficulties with daily chores like bathing, eating, getting dressed, getting in or out of bed, and walking across a room. The chronic conditions include diseases like blood pressure, diabetes, lung disease, heart disease, stroke, cancer, arthritis and psychological disorders. The dichotomous variables created for these conditions are based on whether the individual has been medically diagnosed and whether he has been using prescription drugs or undergoing therapy to treat this specific condition in the previous two years, to get relatively more objective indicators of physical health.

Mental health: The existing literature has mostly overlooked measures of mental health. The studies that do include mental health only account for depression while paying no attention to cognitive disorders. I measure mental health by using information on depression and cognition as well as other diseases like Alzheimer's and Dementia. These mental health conditions are also validated with prescription drug use information. In the HRS, depression is measured in a 0–8 scale, as defined by the Center for Epidemiologic Studies on Depression (CESD). This CESD score measures the sum of adverse mental health symptoms for the past week, based on if the respondent felt depressed, felt that everything was an effort, had restless sleep, was not happy, felt lonely, felt sad, could not get going, and did not enjoy life. Studies have confirmed this to be a valid and reliable indicator for incidence of major depression in older adults (Irwin et al. 1999). Information on measures of cognitive functioning is also included in my analysis. The cognitive functioning measures include immediate and delayed word recall, the serial 7's test, counting backwards, naming tasks (e.g., date-naming), and vocabulary questions. In addition to the individual cognitive functioning measures, the HRS also derives three cognition summary indices. The total recall index which is available for all waves is a concise summary of the immediate and delayed word recall tasks. The mental status index adds the scores from counting, naming, and vocabulary. To maintain consistency across waves, I have used the total cognition score in this study, which sums the total recall and mental status indices and thus ranges from zero to thirty-five.

Other health measures: In addition to self-reported health and measures of physical and mental health, some other measures of health (ignored in existing studies) are also used in this study. These measures include body mass index (BMI), work related stress, physical effort at work, extent of physical exercise, number of nights at hospital, number of doctor visits, risky behaviors like smoking and drinking, and out-of-pocket-medical expenditure.

Descriptives

One way to analyze the impact of health status on the decision to exit full-time employment is to investigate whether or not there are baseline differences (in 1992) in health status between those that are observed working full-time throughout the sample and those that are observed exiting full-time work via one of the routes. Table 1 presents such a comparison for some important standard measures of health. In general, those that subsequently exit from full-time work seem to have worse baseline measures of health than those that remain working full-time. For example, those that exit from full-time employment via complete retirement are more likely to report poor health, ADL difficulties, multiple chronic conditions, depression, and psychological problems. A similar pattern is observed for individuals who subsequently exit via other routes. There is statistically significant difference in means of the health outcome measures for samples of working individuals and individuals who exit via one of the routes as reported in Table 1. The baseline differences in socio-demographic and economic variables for the individuals working full-time across all waves and those that subsequently exit via one of the routes are reported in Table 2. The individuals who subsequently exit via complete retirement route are older, more likely to be male, married and in blue-collar jobs with lower individual and household income as compared to those that remain in labor market full time. For most health measures baseline difference in means for individuals who continue in full-time work and ones who subsequently exit is significantly driven by exit categories-complete retirement and disability. For few health measures, there is statistically significant difference in mean for exit categories partial retirement and part-timework (compared to full-time work) but for unemployment exit route, the differences are not statistically significant. However, mixed results are observed for socio-demographic variables. Baseline difference in a majority of socio-demographic and economic variable means for individuals who

Table 1 Descriptive statistics for some health outcome variables measured in 1992

Variables	No exit/ Working full-time	Individuals working fulltime in 1992 who subsequently exit for the first time via:				
		Complete retirement route	Partial retirement route	Part-time work route	Unemployment route	Disability/Not in labor force route
Health outcomes (Dichotomous)						
Has poor health***	0.01	0.14	0.07	0.03	0.01	0.09
Has good health	0.62	0.31	0.57	0.58	0.66	0.36
Presence of ADL difficulty***	0.01	0.07	0.03	0.01	0.01	0.03
Presence of multiple chronic conditions***	0.56	0.66	0.60	0.55	0.01	0.01
Presence of high BP	0.30	0.35	0.24	0.31	0.33	0.31
Presence of diabetes**	0.06	0.11	0.09	0.08	0.03	0.12
Presence of cancer	0.04	0.04	0.05	0.06	0.01	0.03
Presence of heart Disease**	0.07	0.12	0.12	0.08	0.07	0.09
Presence of lung disease**	0.02	0.04	0.09	0.03	0.03	0.03
Presence of arthritis**	0.29	0.35	0.28	0.27	0.24	0.40
Presence of depression***	0.40	0.43	0.37	0.50	0.50	0.75
Total cognition index score***	24.64	23.59	23.48	23.69	24.31	24.07
Presence of Psychological Problem**	0.03	0.12	0.05	0.05	0.03	0.07
Health outcomes (Non-Dichotomous)						
No. of nights at hospital*	0.39	1.73	0.53	0.38	0.23	0.60
No. of doctor Visits*	3.17	4.42	3.57	2.65	2.83	4.58
Out of pocket medical expenditure*	0.03	0.05	0.04	0.02	0.04	0.10
BMI	26.98	28.01	27.09	27.02	26.90	28.30
Observations	3285	349	148	209	70	67

Sample of individuals at least 50 years old and working full-time in 1992 drawn from the Health and Retirement Study (1992–2010). In the Health Outcomes column, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ indicate the statistical significance of the health measures for the test of equality of means across all samples. For the dichotomous variables, 1 indicates YES and 0 indicates NO

Table 2 Descriptive statistics for socio-demographic variables measured in 1992

Variables	Individuals working fulltime in 1992 who subsequently exit for the first time via:					
	No Exit/ Working full-time	Complete retirement route	Partial retirement route	Part-time work route	Unemployment route	Disability/Not in labor force route
Socio-demographic (Non-Dichotomous)						
Age***	55.67	58.62	59.49	55.93	55.31	54.82
Mother's age	73.56	73.46	73.41	72.96	70.33	70.54
Father's age	70.61	70.00	69.76	69.06	68.09	69.91
Years of schooling**	12.84	12.32	12.80	11.93	12.70	11.43
Individual income***	31,605.73	28,862.54	31,604.66	19,488.27	27,437.91	18,031.25
Household income***	54,537.24	51,095.82	57,483.95	51,897.82	46,381.00	35,770.48
Socio-demographic (Dichotomous)						
Female***	0.43	0.36	0.28	0.66	0.44	0.72
Black**	0.14	0.15	0.09	0.21	0.20	0.22
Hispanic	0.02	0.01	0.03	0.02	0.01	0.01
Other race	0.06	0.05	0.05	0.11	0.11	0.16
Married**	0.75	0.77	0.84	0.72	0.64	0.58
Blue collar worker	0.28	0.33	0.28	0.23	0.36	0.39
Has religious preference	0.95	0.96	0.97	0.96	0.91	0.93
Native	0.91	0.96	0.96	0.87	0.86	0.87
Father's education (>8 yrs)	0.57	0.54	0.55	0.55	0.74	0.51
Mother's education (>8 yrs)	0.61	0.60	0.63	0.58	0.71	0.55
Has govt. insurance***	0.05	0.11	0.11	0.03	0.03	0.10
Has private insurance***	0.14	0.13	0.20	0.21	0.09	0.12
Has employer insurance***	0.87	0.86	0.77	0.66	0.76	0.70

Table 2 (continued)

Variables	Individuals working fulltime in 1992 who subsequently exit for the first time via:					
	No Exit/ Working full-time	Complete retirement route	Partial retirement route	Part-time work route	Unemployment route	Disability/Not in labor force route
Has non-employer Insurance***	0.06	0.08	0.11	0.15	0.07	0.13
Has no insurance***	0.07	0.06	0.11	0.19	0.17	0.16
Has pension DC***	0.20	0.16	0.16	0.12	0.23	0.12
Has pension DB***	0.27	0.34	0.22	0.14	0.16	0.27
Has pension Both***	0.18	0.19	0.16	0.06	0.07	0.07
Don't Know about pension**	0.01	0.01	0.01	0.01	0.01	0.01
Observations	3285	349	148	209	70	67

Sample of individuals at least 50 years old and working full-time in 1992 drawn from the Health and Retirement Study (1992–2010). In the Health Outcomes column, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$ indicate the statistical significance of the health outcome variables for the test of equality of means across all samples. For the dichotomous variables, 1 indicates YES and 0 indicates NO continue in full-time work and ones who subsequently exit are driven by exit categories part-time work and disability.

Health Indices

To mitigate the potential difficulties arising due to use of subjective self-reported measure of health, Bound (1991, Bound et al. 1998) suggested an approach that involves estimating a model of self-reported health as a function of more objective measures of health to define a latent 'health stock' variable. This health stock variable is then used as a measure of health in a model of retirement / exit from full-time employment. This idea of constructing a health stock is in many ways analogous to using objective measures of health to instrument for the endogenous and potentially error-ridden self-reported health variable. In this paper, the latent health stock variable is predicted by using an ordered probit model for self-reported health, where the ordered measure of self-reported health is regressed on a set of relatively more objective health measures reported in Table 1. More health problems are associated with a lower order of the latent health stock. Unfortunately, this method for creating a latent health stock is not very effective at suggesting how the different individual health measures are weighted. This can be a problem because clearly neither high blood pressure nor diabetes is the same as cancer. Physical health outcomes are also clearly different from mental health outcomes. On the other hand, including every health measure separately in a regression model will make it cumbersome. Hence principal component analysis is used as a comprehensive way to extract eight[10] meaningful factors (indices) from twenty-eight individual health outcomes. For each factor, it is possible to note how the different individual health measures weigh. [11] It is important to note that self-reported health does not load heavily in any factor, which implies that it is not the best indicator of health of an individual. Based on the health outcomes that load heavily in each factor[12] I label them: Factor 1: Has chronic conditions, Factor 2: Has functional limitation, Factor 3: Hospital stay, Factor 4: Has cognitive functioning problems, Factor 5: Has depression, Factor 6: Lack of physical exercise, Factor 7: Has cancer, and Factor 8: Has lifestyle behavioral problems.

Empirical Method

In this paper, a standard proportional hazard model is used to estimate the impact of health on the duration of full-time employment, where time is measured in two-year waves. In some specifications, only initial employment spells are

[10] Factors having eigen-values greater than 1 are retained. These eight factors are the "principal components" of the health of individuals in the sample. In other words, they represent perceived health status of individuals in the sample in the best possible objective way. From twenty eight diverse health outcome variables. Eight factors with Eigen value greater than 1 are generated using Principal Component Factor Analysis which is used to create the health indices used as explanatory variable in the hazard model.

[11] Principal Component Analysis results are reported in Appendix Tables 10 and 11.

[12] Refer to Appendix Table 11 to see the factor loadings (i.e. which measures load heavily in each factor).

included in the sample, while in other specifications I include subsequent employment spells as well. Another way in which I differentiate the model is to combine all five exit routes in some specifications (a combined risk or lumped risk model) while other specifications each exit route is treated separately (a competing risk model).

More formally, the competing risk proportional hazard model is given by:

$$H_j\ (t) = H_0\ (t)^*\ \exp\left(X'_{it}\beta\right) \qquad (1)$$

Here, j is an index for each of the five exit routes and Xit is the vector of covariates that vary with time while $H_0(t)$ is the baseline hazard that only depends on time but not individual covariates which means it is common for all units. The impact of the observable characteristics is parametrically estimated using the standard proportional hazard functional form exp. $(X'_{it}\beta)$.

Given that the hazard is not constant over time (time-dependency of hazard rates), it is important to choose a suitable parametric distribution for estimating the baseline hazard. If the chosen distribution correctly characterizes the time-dependency, then the parameter estimates are likely to be more precise than the parameter estimates of semi parametric or non-parametric models where the time-dependency is left unspecified. Hence, there are advantages of using a suitable parametric model. But problems may arise if a wrong parametric form is chosen. The most common approach for choosing an appropriate parametric model is based on using the Akaike Information Criterion (AIC). It is based on penalizing the log likelihood to reflect the number of parameters being estimated by different models (distributions) and comparing them. Although the best fitting distribution is the one with the largest log likelihood, the one with smallest AIC is most preferred. Table 3 presents the log likelihood and AIC information for different parametric models. Given the smallest AIC, the Weibull distribution is chosen for parametrically estimating the baseline hazard.[13]According to the proportional hazard specification stated earlier, the Weibull hazard rate is given as:

$$H\ (t, X) = \lambda\ p\ (\lambda t)\ p{-}1 \qquad (2)$$

Where, $\lambda_i = e^{Xi\beta}$ and p is the shape parameter.

In all specifications, in addition to the socio-demographic and economic variables reported in Table 2, each specification I estimate includes spousal health and work status, occupations, census regions, expected longevity[14] and controls for general economic conditions (through wave dummies). In order to estimate the model with standard software, an independence assumption across the exit routes is imposed.

[13] I have also estimated the following other parametric models Gompertz (proportional hazard model), Log normal, Log logistic and Gamma (Accelerated Failure Time models) for all specifications (not reported) and found the time ratios (similar to hazard ratios). In the generalized gamma model, the Wald test for κ =1 provides support for adopting the Weibull distribution.

[14] Expected subjective probability of living until age 85.

Table 3 Log likelihood and akaike information criterion for different parametric models

DISTRIBUTION	ALL RISKS: All Routes		COMPETING RISKS									
			Exit Route 1: Complete Retirement		Exit Route 2: Partial Retirement		Exit Route 3: Part-time Work		Exit Route 4: Unemployment		Exit Route 5: Disability and Not in Labor force	
	Log-likelihood	AIC	Log-likelihood	AIC	Log-likelihood	AIC	Log-likelihood	AIC	Log-likelihood	AIC	Log-likelihood	AIC
Exponential	-4441.75	8957.49	-2783.79	5641.58	-1786.91	3647.82	-1538.07	3150.14	-616.79	1307.59	-640.41	1354.82
Weibull	-4182.89	7871.98	-2663.72	5185.80	-1747.70	3402.30	-1487.03	2985.92	-591.56	1254.86	-591.30	1253.54
Log normal	-3897.99	8014.04	-2544.90	5233.87	-1663.15	3453.57	-1454.96	3027.67	-589.43	1258.58	-589.49	1255.00
Log logistic	-3969.01	8441.77	-2578.93	5403.45	-1688.78	3571.40	-1475.84	3050.06	-591.28	1259.12	-589.56	1258.61
Generalized Gamma	-3914.14	7997.98	-2547.44	5192.88	-1671.37	3440.73	-1465.74	2999.28	-589.97	1256.73	-591.23	1256.43

The table shows the log likelihood and AIC for the different parametric models. The log normal has the highest log likelihood across all exit routes and should be the best fitting model, but the Weibull has the lowest AIC, hence is the most preferred one and used for estimating the baseline hazard in all subsequent regression specifications

Then this independence assumption is tested by estimating Martingale residuals for each exit route and checking their correlations for statistical significance, as in Borgan and Langholz (2007) and Marton et al. (2010).

Hazard models may be plagued by duration dependence, which arises due to unobserved heterogeneity.[15] Ignoring unobserved heterogeneity may exaggerate the rate of failure for some individuals and underestimate the rate of failure for others. In this context, unobserved heterogeneity is addressed through the addition of an additional random parameter or "frailty term" to the model. In the proportional hazards model, the hazard rate increases or decreases with the covariates. The problem is that if there are unmeasured or unobserved 'frailties', then the hazard rate will be a function of the covariates and the frailties. Eq. (1) may be rewritten as:

$$Hj\ (t) = H_0\ (t)^{*}\exp\left(X'_{it}\beta + \varepsilon_i\right) \tag{3}$$

Or,

$$Hj\ (t) = H_0.v_i.\exp\left(X'_{it}\beta\right) \tag{4}$$

Where $v_i = \exp.(\varepsilon_i)$.

So, the frailty term acts multiplicatively on the hazard rate. The hazard rate is conditional on both the covariates and the frailty. For identification purposes, it is assumed that mean of v = 1 and the variance is equal to some unknown finite parameter θ. Then the unobserved heterogeneity or frailty is modeled using a gamma distribution and effectively the frailty variance θ is estimated. The hypothesis that θ = 0 may be tested using a likelihood ratio test to determine whether unobserved heterogeneity is something to worry about in the model.

This is equivalent to the inclusion of a random effects term in a standard panel data model. In some specifications, I include a frailty term that is modeled parametrically using the gamma distribution.[16]There is no hazard model equivalent to a fixed effect

[15] Individuals with the same observed characteristics are not identical. The notion of unobserved heterogeneity amounts to observations being conditionally different (heterogeneous) in terms of their hazards in ways that are unaccounted for in the standard hazard model. In other words, some observations are more "frail" than others.

[16] Competing risk models that include controls for unobserved heterogeneity and allow for correlation in the unobserved heterogeneity across exit routes (semiparametric estimation) are presented in Butler et al. (1989) and Canals-Cerdá and Gurmu (2007).

panel data model. In the results section below, I investigate this potential limitation by estimating standard linear probability models of retirement with both fixed and random effects to see if there is a big different in the coefficients.[17]

Results

Controlling for Unobserved Heterogeneity in Hazard Models Vs. Panel Data Models

I start by estimating a simple linear probability model of complete retirement using the standard health measures from existing studies and only the first five waves of the HRS.[18] Both fixed effect (FE) and random effect (RE) models are presented in panel A of Table 4. These results suggest poor self-reported health, multiple ADL difficulties and chronic diseases, heart disease and stroke lead to an increase in probability of complete retirement. In this case, the Hausman-Wu test rejects the null hypothesis that the difference in the coefficients in the FE and RE models is not systematic. This implies that the FE model does a better job of controlling for unobserved heterogeneity. Panel B illustrates that adding an additional five waves to the sample does not affect the coefficient estimates in a major way, except that having psychological problem now leads to a statistically significant increase in probability of complete retirement. The Hausman-Wu test again rejects the hypothesis that the RE and FE coefficients are the same.

As mentioned in Section 3.4, there are advantages associated with using the health indices (factors) that come from a PCA to control for health status, rather than the limited individual health measures typically used in the literature. Panel C includes the eight health indices described in Section 3.4 rather than the typical health indicators from the literature.[19] When we use 10 waves of the HRS and include our PCA health indices to measure individual health status, the Hausman-Wu test cannot reject the hypothesis of equality of the RE and FE coefficients. Given that the RE and FE coefficients are so similar in the linear probability model framework, I have confidence that the lack of a fixed effects specification within a hazard model framework is not a serious limitation in my subsequent analysis presented below. In other words, this implies that the individual random effects (frailty) in my hazard model will control for the same unobserved factors as would individual fixed effects.

[17] The drawback of such fixed effects model is that it does not differentiate between full-time employment spells of different duration.

[18] Some of these studies are Dwyer and Mitchell (1999), McGarry (2004) and Miah and Wilcox-Gök (2007).

[19] In both panels, B and C chronic conditions lead to a statistically significant increase in probability of complete retirement. However, from the estimates in panel B it is only possible to say whether the presence of multiple chronic conditions (dichotomous) affect the probability of complete retirement. However, panel C has a more objective measure of chronic illness because Factor 1 includes information on the count of chronic conditions based on medical diagnosis as well as information on intake of prescriptions drugs specific for treating those chronic conditions. Similarly, in Panel B it is possible to state that presence of psychological disorder (dichotomous) raises the probability of complete retirement. While in Panel C, factor 5 indicates a high score on CESD scale (depression), having work-related stress and medically diagnosed psychological disorder for which the individual also takes prescription medication, increases the probability of complete retirement. Therefore, the panel C health factors are clearly more objective than the panel B standard health measures.

Table 4 Fixed and random effect estimates in linear model

	(A) Using HRS Waves (1992–2000)		(B) Using HRS Waves (1992–2010)		(C) Using HRS Waves (1992–2010)	
	FE	RE	FE	RE	FE	RE
Health outcomes						
Self-Reported Poor	0.053***	0.061***	0.088***	0.093***		
Health	(0.015)	(0.014)	(0.011)	(0.010)		
Has Activity of Daily	0.030***	0.024***	0.040***	0.051***		
Living Difficulty	(0.011)	(0.007)	(0.012)	(0.012)		
Has Multiple Chronic	0.070***	0.040***	0.030***	0.015**		
Conditions	(0.017)	(0.011)	(0.007)	(0.007)		
Has High Blood Pressure	0.013	0.010	0.026	0.033		
	(0.012)	(0.007)	(0.007)	(0.005)		
Has Diabetes	0.021	0.027	0.011	0.026		
	(0.016)	(0.009)	(0.008)	(0.007)		
Has Heart Disease	0.043***	0.036***	0.026***	0.038***		
	(0.014)	(0.009)	(0.008)	(0.007)		
Has Lung Disease	0.036	0.025	0.054**	0.044**		
	(0.023)	(0.014)	(0.012)	(0.009)		
Had Stroke	0.086***	0.028	0.060***	0.050***		
	(0.027)	(0.018)	(0.014)	(0.012)		
Has Cancer	0.023	0.001	0.027*	0.039*		
	(0.012)	(0.009)	(0.010)	(0.007)		
Has Arthritis	0.013	0.021	0.017	0.024		
	(0.018)	(0.017)	(0.007)	(0.005)		
Has Psychological	0.005	0.014	0.029**	0.019**		
Problem	(0.020)	(0.013)	(0.012)	(0.009)		
Health indices						
Factor 1: Has Chronic Conditions					0.073***	0.072***
					(0.004)	(0.003)
Factor 2: Has Functional Limitations					0.100***	0.102***
					(0.003)	(0.003)
Factor 3: Hospital Stay					0.059	0.052
					(0.004)	(0.002)
Factor 4: Has Cognitive Disorders					0.121	0.119
					(0.003	(0.003)
Factor 5: Has Depression					0.062***	0.065***
					(0.003)	(0.002)
Factor 6: Lack of Physical Exercise					0.016***	0.024***
					(0.002)	(0.002)
Factor 7: Has Cancer					0.015*	0.019*
					(0.004)	(0.003)
Factor 8: Has Lifestyle					0.021	0.011

Table 4 (continued)

	(A) Using HRS Waves (1992–2000)		(B) Using HRS Waves (1992–2010)		(C) Using HRS Waves (1992–2010)	
Behavioral Problems					(0.009)	(0.007)
OBSERVATIONS	18,428	18,428	32,707	32,707	32,707	32,707

This table shows the FE and RE regression results for effect of health on complete retirement for individuals over 50 years of age and employed full-time in 1992. Panel A uses the first five waves of the HRS (similar to exiting studies). Panel B uses another five additional waves. Both Panels A and B use the standard physical and mental dichotomous health outcome measures used in existing studies. Panel C uses ten waves of the HRS and the health indices (factors) created by PCA from twenty-eight health outcome variables. All panels control for the socio-demographic variables reported in Table 2. Also controlling occupations, the wave dummies, census regions, expected mortality and work and health status of spouse, which are not been reported here

The standard errors are reported in parentheses and, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$

Hazard Model

The hazard ratios from parametric hazard model (Weibull) are reported for initial exit and multiple exit models with and without frailty for combined risk competing risks specification. Table 5 reports the hazard ratios for the health indices (factors) for both the combined risk and competing risks specifications estimated without a frailty term over all initial employment spells. The latent health stock variable has statistically significant hazard ratios 0.845 and 0.796 for the combined risk model and the complete retirement exit route, respectively. This implies an increase in one's latent health stock (i.e. better self-reported health) makes an individual 15.5% less likely to exit from full-time employment in general and 20.4% less likely to exit via complete retirement route. The functional limitations factor has statistically significant hazard ratios of 1.101, 1.136 and 1.278 (significant at 1% level) for the combined risk model, the complete retirement exit route, and the disability/not in labor-force exit route, respectively. This implies that multiple functional limitations (ADL difficulties) that limit mobility and work, increase the probability of exit from full-time employment by 10.1% in general, 13.6% via the complete retirement exit route, and 27.8% via the disability exit route. More generally, a hazard ratio for an independent variable greater than 1 implies that the presence of (or an increase in) that variable leads to an increase in the likelihood of an exit (i.e. worse chances of survival in full-time employment). The opposite is true if the estimated hazard ratio is less than 1 (i.e. better chances of survival in full-time employment). The magnitude of change is calculated as (1-Hazard Ratio). A positive value would signify better chances of survival (lower likelihood of exit) while a negative value would signify worse chances of survival (higher likelihood of exit).The p-values are for the hypothesis test that the hazard ratio for the variable in question is equal to 1 (i.e. no effect). The magnitudes of these effects can be difficult to interpret, because they are relative probabilities. Therefore, the absolute effect[20]associated with each independent variable area also reported. For each exit route, these absolute effects can be compared to the average probability of exiting via that route.[21] For example, the

[20] Here Absolute Effect = Hazard Ratio * Average Exit Probability via that route. Average Exit Probability = No. of Exits / (No. of Spells * Average Spell Length).

[21] The average exit probabilities are reported at the bottom of Tables 5, 6, and 7, 7 and 8.

Table 5 Competing risk model (weibull distribution) for initial exit from full-time employment (without frailty)

Health outcomes	All risks		Competing risks									
			Exit route 1:Complete retirement		Exit route 2:Partial retirement		Exit route 3:Part-time work		Exit route 4:Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
Latent health stock: Self-reported health	0.845***	13.89%	0.796***	6.71%	1.112	4.29%	0.971	2.46%	0.965	0.71%	0.969	0.82%
	(0.033)		(0.043)		(0.103)		(0.147)		(0.095)		(0.175)	
Factor 1: Chronic conditions	0.988	16.24%	1.013	8.54%	0.986	3.81%	0.982	2.49%	0.827	0.61%	0.922	0.78%
	(0.023)		(0.033)		(0.050)		(0.056)		(0.088)		(0.077)	
Factor 2: Functional limitation	1.101***	18.10%	1.136***	9.58%	1.070	4.13%	0.932	2.36%	1.158	0.85%	1.278***	1.08%
	(0.021)		(0.029)		(0.050)		(0.056)		(0.099)		(0.073)	
Factor 3: Hospital stay	1.016	16.70%	1.012	8.54%	0.924	3.56%	1.076	2.73%	1.078	0.79%	1.133	0.96%
	(0.022)		(0.030)		(0.048)		(0.056)		(0.105)		(0.086)	
Factor 4: Cognitive functioning	1.075***	17.67%	1.086**	9.16%	0.983	3.79%	1.062	2.70%	0.880	0.65%	1.321***	1.12%
	(0.030)		(0.043)		(0.059)		(0.070)		(0.102)		(0.140)	
Factor 5: Depression	1.088***	17.88%	1.088***	9.18%	0.910*	3.51%	1.091*	2.77%	1.186*	0.87%	1.291***	1.09%
	(0.023)		(0.033)		(0.047)		(0.054)		(0.106)		(0.084)	
Factor 6: Physical exercise	0.993	16.32%	0.971	8.19%	1.081*	4.17%	1.025	2.60%	0.823**	0.61%	1.051	0.89%
	(0.021)		(0.028)		(0.047)		(0.053)		(0.079)		(0.085)	
Factor 7: Cancer factor 8: Life style0.938***	15.42%	0.924***	7.80%	0.952	3.67%	0.940	2.39%	0.797*	0.59%	0.934	0.79%	
	(0.020)		(0.028)		(0.043)		(0.052)		(0.097)		(0.077)	
Factor 8: Life style	1.104***	18.15%	1.144***	9.65%	1.137***	4.38%	1.037	2.63%	1.042	0.77%	1.063	0.90%

Table 5 (continued)

Health outcomes	All risks		Competing risks									
			Exit route 1:Complete retirement		Exit route 2:Partial retirement		Exit route 3:Part-time work		Exit route 4:Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
	(0.024)		(0.036)		(0.053)		(0.055)		(0.101)		(0.095)	
# Spells	4128		4128		4128		4128		4128		4128	
# Exits	3272		1680		769		505		148		170	
Avg. spell length	4.82		4.82		4.82		4.82		4.82		4.82	
Avg. exit probability	16.44%		8.44%		3.86%		2.54%		0.74%		0.85%	

Table 5 uses data from the Health and Retirement Study (1992–2010) for a proportional hazard model in competing risks framework for individuals who were at least 50 years old and worked full-time in 1992. The underlying baseline hazard is parametrically estimated using Weibull distribution where the shape parameter p is approximately close to 2. The hazard ratios and absolute effect for the single time exit scenario are reported. Absolute Effect = Hazard Ratio * Average Exit Probability via that route. Average exit probability = no. of exits/(no. of spells * average spell length). The dependent variable is duration of full-time employment while the independent variables are health outcome indices constructed by principal component factor analysis. The competing risks in this framework are the different routes of exit from full-time employment. The combined risk model has also been reported in panel A in both tables. Controlling for the socio-demographic variables reported in Table 2. Also controlling occupations, the wave dummies, census regions, expected mortality and work and health status of spouse, which have not been reported here. This specification does not control for frailty (unobserved heterogeneity). In both tables standard errors are reported in parentheses and *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

absolute effect associated with the functional limitations factor in the combined risk model suggests that an increase in functional limitations makes an individual 18.10% more likely to exit from full-time employment, which is greater than the average exit probability of 16.44% for every period. An increase in functional limitations also makes an individual 9.58% and 1.08% more likely to exit via complete retirement and disability routes respectively, which is higher than the average probabilities to exit via those routes (8.44% and 0.85%, respectively) every wave. Other statistically significant health factors include depression, risky lifestyle behavior like drinking and smoking and cancer. It is observed that an increase in the cancer factors leads to a lower probability of exit via any route in general and the complete retirement route in particular.

In Table 6, I present the same models but estimate them using both initial and subsequent employment spells (multiple exit model). While this increases the sample size, it does not generate large changes in the coefficient estimates. Higher latent health stock (better self-reported health) still makes an individual less likely to exit. For physical health conditions, an increase in functional limitation factor makes an individual more likely to exit via any route in general and via the complete retirement and disability routes in particular. Among the mental health factors, an increase in depression raises the probability of exit via any route in general and through complete retirement route in particular. It is interesting to note that cognitive functioning disorders factor have no statistically significant effect on exit through the complete retirement route, but an increase in problems related to cognitive functioning makes an individual 1.12% more likely to exit via the disability route. Increase in risky behaviors makes an individual 9.47% more likely while cancer makes one 7.78% less likely to exit via the complete retirement route.

In Table 7, I re-estimate the models from Table 6 on the same sample, but now include a frailty term in the model to control for unobserved heterogeneity (this represents the most complete model).[22] These results suggest that increase in latent health stock makes an individual 13.72% less likely to exit from full-time employment in general, 6.44% less likely to exit via complete retirement and 0.37% (significant at 1% level) less likely to exit via disability route. These are slightly lower than the average probabilities associated with these exit routes. For the indices of physical health, an increase in the functional limitations factor increases the likelihood of exit from full-time work by 18.88% overall, by 9.50% for the complete retirement exit route and by 1.07% for the disability exit route (significant at 1% level). These are higher than the average exit probabilities associated with these exit routes (bottom of Table 7). For mental health factors, an increase in depression factor increases the likelihood of exit from full-time employment for an individual by 18.68% in general, 9.06% via the complete retirement route (significant at 1% level), 0.90% via unemployment route (significant at 5% level) and 3.04% via part-time work route (significant at 10% level). Increases in cognitive problems factor have no statistically significant effect on the likelihood of exit via complete retirement, but increases the likelihood of exit via the disability exit route by 1.14% (significant at 1%). The risky behavior factor leads to 9.59% (significant at 1% level) and 4.82% (significant at 5% level) higher probability of exit via complete retirement and partial retirement respectively, while the cancer

[22] The initial exit model with frailty is presented in Appendix Table 12.

Table 6 Competing risk model (weibull distribution) for exiting full-time employment allowing for multiple spells of employment (without frailty)

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
Latent health stock: Self-reported health	0.848***	14.36%	0.785***	6.56%	1.101	4.72%	0.978	2.68%	0.968	0.67%	0.568***	0.48%
	(0.030)		(0.040)		(0.090)		(0.090)		(0.088)		(0.063)	
Factor 1: Chronic conditions	1.010	17.10%	1.036	8.65%	1.027	4.40%	0.982	2.69%	0.846	0.59%	0.944	0.80%
	(0.022)		(0.032)		(0.046)		(0.052)		(0.087)		(0.077)	
Factor 2: Functional limitation	1.086***	18.39%	1.128***	9.42%	1.037	4.48%	0.932	2.55%	1.133	0.79%	1.257***	1.07%
	(0.020)		(0.028)		(0.044)		(0.051)		(0.093)		(0.071)	
Factor 3: Hospital stay	1.010	17.10%	1.017	8.49%	0.917	3.93%	1.069	2.93%	1.005	0.70%	1.145*	0.97%
	(0.020)		(0.028)		(0.042)		(0.050)		(0.100)		(0.0841)	
Factor 4: Cognitive functioning	1.067**	18.08%	1.079	9.01%	0.967	4.14%	1.057	2.89%	0.903	0.63%	1.322***	1.12%
	(0.028)		(0.041)		(0.052)		(0.065)		(0.102)		(0.138)	
Factor 5: depression	1.078***	18.26%	1.078***	9.00%	0.936	4.01%	1.060	2.90%	1.220**	0.85%	1.290	1.10%
	(0.021)		(0.031)		(0.043)		(0.049)		(0.101)		(0.083)	
Factor 6: Physical exercise	1.001	16.95%	0.965	8.06%	1.099**	4.32%	1.025	2.81%	0.824**	0.58%	1.077	0.92%
	(0.019)		(0.027)		(0.043)		(0.049)		(0.077)		(0.085)	
Factor 7: Cancer	0.951***	16.10%	0.932**	7.78%	0.966	4.14%	0.967	2.65%	0.896	0.62%	0.917	0.78%
	(0.019)		(0.0265)		(0.038)		(0.047)		(0.095)		(0.075)	
Factor 8: Life style	1.089***	18.45%	1.134***	9.47%	1.096**	4.70%	1.032	2.83%	1.085	0.76%	1.026	0.87%

Table 6 (continued)

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
	(0.022)		(0.034)		(0.046)		(0.051)		(0.103)		(0.089)	
# Spells	4868		4868		4868		4868		4868		4868	
# Exits	3817		1883		968		617		157		192	
Avg. spell length	4.63		4.63		4.63		4.63		4.63		4.63	
Avg. exit probability	16.94%		8.35%		4.29%		2.74%		0.70%		0.85%	

Table 6 uses data from the Health and Retirement Study (1992–2010) for a proportional hazard model in competing risks framework for individuals who were at least 50 years old and worked full-time in 1992. The underlying baseline hazard is parametrically estimated using Weibull distribution where the shape parameter p is approximately close to 2. It reports hazard ratios and absolute effect for the multiple exits scenario. Absolute Effect = Hazard Ratio * Average Exit Probability via that route. Average exit probability = no. of exits/ (no. of spells * average spell length). The dependent variable is duration of full-time employment while the independent variables are health outcome indices constructed by principal component factor analysis. The competing risks in this framework are the different routes of exit from full-time employment. The combined risk model has also been reported in panel A in both tables. Controlling for the socio-demographic variables reported in Table 2. Also controlling occupations, the wave dummies, census regions, expected mortality and work and health status of spouse, which have not been reported here. This specification does not control for frailty (unobserved heterogeneity). In both tables standard errors are reported in parentheses and *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 7 Competing risk model (weibull distribution) for exiting full-time employment allowing for multiple spells of employment (with frailty)

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
Latent health stock: Self-reported health	0.810***	13.72%	0.772***	6.44%	1.090	4.68%	0.947	2.59%	0.916	0.64%	0.430***	0.37%
	(0.040)		(0.044)		(0.109)		(0.115)		(0.192)		(0.077)	
Factor 1: Chronic conditions	0.985	16.69%	1.027	8.58%	0.990	4.25%	0.942	2.58%	0.845	0.59%	0.962	0.82%
	(0.028)		(0.034)		(0.053)		(0.065)		(0.107)		(0.0890)	
Factor 2: Functional limitation	1.115***	18.88%	1.138***	9.50%	1.049	4.51%	0.934	2.56%	1.177	0.82%	1.255***	1.07%
	(0.026)		(0.030)		(0.051)		(0.061)		(0.126)		(0.076)	
Factor 3: Hospital stay	1.011	17.12%	1.015	8.48%	0.929	3.99%	1.079	2.95%	0.984	0.69%	1.151*	0.98%
	(0.026)		(0.030)		(0.050)		(0.064)		(0.117)		(0.092)	
Factor 4: Cognitive functioning	1.073**	18.17%	1.085	9.06%	0.973	4.18%	1.089	2.98%	0.852	0.59%	1.336***	1.14%
	(0.036)		(0.045)		(0.062)		(0.087)		(0.121)		(0.150)	
Factor 5: Depression	1.103***	18.68%	1.085***	9.06%	0.947	4.07%	1.110*	3.04%	1.289**	0.90%	1.319	1.12%
	(0.028)		(0.034)		(0.050)		(0.066)		(0.133)		(0.099)	
Factor 6: Physical exercise	0.972	16.47%	0.956	7.99%	1.090*	4.67%	1.017	2.78%	0.769**	0.53%	1.063	0.91%
	(0.025)		(0.029)		(0.052)		(0.062)		(0.087)		(0.089)	
Factor 7: Cancer	0.932***	15.78%	0.924**	7.71%	0.951	4.08%	0.920	2.52%	0.888	0.62%	0.904	0.77%
	(0.023)		(0.028)		(0.045)		(0.0592)		(0.108)		(0.080)	
Factor 8: Life style	1.131***	19.15%	1.148***	9.59%	1.124**	4.82%	1.074	2.94%	1.065	0.74%	1.021	0.87%
	(0.032)		(0.038)		(0.059)		(0.070)		(0.124)		(0.0932)	

Table 7 (continued)

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
# Spells	4868		4868		4868		4868		4868		4868	
# Exits	3817		1883		968		617		157		192	
Avg. Spell length	4.63		4.63		4.63		4.63		4.63		4.63	
Avg. Exit probability	16.94%		8.35%		4.29%		2.74%		0.70%		0.85%	

Table 7 uses data from the Health and Retirement Study (1992–2010) for a proportional hazard model in competing risks framework for individuals who were at least 50 years old and worked full- time in 1992. The baseline hazard is estimate using the Weibull distribution where the shape parameter p is approximately close to 2. It reports hazard ratios and absolute effect for the multiple exits scenario. Absolute Effect = Hazard Ratio * Average Exit Probability via that route, where average exit probability = no. of exits/ (no. of spells * average spell length). The dependent variable is duration of full-time employment while the independent variables are health outcome indices constructed by principal component factor analysis. The competing risks in this framework are the different routes of exit from full-time employment. The combined risk model has also been reported in panel A. Controlling for the socio-demographic variables reported in Table 2. Also controlling occupations, the wave dummies, census regions, expected mortality and work and health status of spouse, which has not been reported here. Further, it controls for frailty (unobserved heterogeneity). The standard errors are reported in parentheses and *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Likelihood ratio test rejects the null hypothesis of No Frailty such that there is statistically significant level of unobserved heterogeneity in the data. The shared frailty is estimated assuming an underlying gamma distribution factor leads 7.71% lower likelihood of exit via complete retirement. The likelihood ratio test for the estimates reported in Table 7 rejects the null hypothesis of "no frailty", which implies the existence of unobserved heterogeneity that needs to be accounted for.

The results from the parametric model are compared to the semi parametric Cox proportional hazard model estimates to check for the consistency of the estimates.[23] The hazard ratios from the Weibull parametric model are directly comparable to the hazard ratios reported in the Cox model. The hazard ratios obtained in the Cox proportional hazard model, are qualitatively similar, consistent in statistical significance but slightly smaller in magnitude for the different health indices for both combined risk and competing risks case. The inference drawn about the effect of the health factors on probability of exit from full-time employment for both Weibull and Cox models is similar. Among physical health factors, functional limitation leads to higher likelihood of exit via complete retirement and disability, with the magnitude of the effect being much smaller for the disability route. While for mental health factors, depression leads to higher likelihood of exit via complete retirement and unemployment while cognitive problems lead to higher probability of exit via the disability route. Lifestyle risky behavior also increases the likelihood of exit via complete retirement. In the Cox proportional hazard model, unlike the parametric Weibull model, cancer does not statistically significantly decrease the probability of exit from full-time employment via complete retirement route.

As a robustness check, I examine the impact of changes in health outcomes between waves, rather than simply looking at levels[24] for the overall sample (i.e. without splitting by age). Changes in self-reported overall health, counts of chronic conditions, counts of functional limitations (ADL difficulties), and the onset of memory related diseases between waves are considered.[25] Extreme reductions in these measures between waves serve as a proxy for exogenous changes in health. The results suggest that a major reduction in overall self-reported health increases the likelihood of exit from full-time employment via complete retirement, and disability. Increases in counts of chronic conditions and onset of memory related diseases, between waves increase the likelihood of exit via the complete retirement and the disability exit routes. While increases in functional limitations between waves have no statistically significant effect on exiting via complete retirement, but increases the likelihood of exit via disability.

In summary, the overall sample results indicate that physical health problems (functional limitations) lead to increases in the likelihood of exit from full-time employment in general, which one can attribute to the increase in the likelihood of exit via the complete retirement route and the disability route. The magnitude of the effect is much smaller for the disability route. As for mental health problems, depression increases the likelihood of exit via complete retirement, while cognitive disorders increase the likelihood of exit via the disability exit route (with no statistically significant effect on the likelihood of exit via complete retirement).

[23] Cox-Proportional Hazard Model for multiple exits is reported in Appendix Table 13.

[24] These results are reported in Appendix Table 14.

[25] I define a reduction in overall self-reported health by flagging anyone that went from excellent to poor self-reported health between waves. Similarly, I defined a reductions in health associated with ADL difficulties and chronic conditions by flagging anyone that went from reporting a 0 to a 5 between waves. Finally, anyone with an onset of a memory-related disease was flagged.

Conclusion

This study contributes to the existing literature by empirically modeling the duration of full-time employment of older Americans using a long panel from the Health and Retirement Study. I distinguish between the different exit routes from full time employment and allow for multiple employment spells. Moreover, this study addresses the inherent problem of the subjectivity of health measures in surveys by constructing relatively objective comprehensive indices of physical and mental health that take into account a wide variety of health indicators based on both medical diagnosis and medication. The PCA method used for construction of the health factors (indices) is not only an effective method of data reduction but also helps to get uncorrelated explanatory variables (health factors). This is particularly important because physical and mental health outcomes are likely to be highly correlated which can lead to endogeneity problem and hence biased estimates. The PCA analysis helps to address this issue although the causal effect of the constructed health factors is not strongly established. Moreover, unlike existing studies, I am able to distinguish between different dimensions of physical and mental health (functional limitations versus chronic conditions and depression versus cognition) and their impact on continued employment.

Consistent with the findings of most existing studies, the main inferences drawn from the results of the most complete model (multiple spells with frailty) indicate that better self-reported health decreases the likelihood of exit from full- time employment. It is also found that physical and mental health problems are both impediments to continued work. Few studies that have explored similar research question based on HRS are divided in their findings. Dwyer and Mitchell (1999) find evidence that overall poor health leads to early retirement plan among men. But few other studies like McGarry (2004) and Miah and Wilcox-Gök (2007) find opposite results. McGarry (2004) finds that poor health has very strong correlation with the decision to remain employed. Similarly, Miah and Wilcox-Gök (2007) find that chronic illness is associated with higher probability of retiring later because individuals with chronic illnesses are able to accumulate less assets over time. Given the lack of consensus in existing literature it is useful to explore this association between health and retirement. The findings of this paper are more informative. I am able to disaggregate the wide variety of health outcomes and observe their differential impact on duration of full-time work. Among physical health factors while functional limitations lead to a higher likelihood of exit from full-time employment, incidence of cancer leads to a decrease in likelihood of exit from full-time work. Among mental health factors, depression leads to an increased likelihood of exit via complete retirement while cognitive problems have no statistically significant impact. Moreover, due to the competing risks specification in hazard analysis framework I am able to distinguish between different routes of exit i.e. some health conditions may decrease the likelihood of continuing in full-time employment but need not imply full-retirement i.e. due to certain health problems like depression an individual may exit from full-time work into part-time work or unemployment while cognitive disorders leads to an increase in likelihood of exit from full-time employments via disability route.

These results produce targets for policies that seek to improve the health of older working Americans. Improving the health of older workers means they can be retained in the labor force for an extended period of time, which would result in decreased training costs for replacement workers, the ability to maintain the experience and productivity of these older workers, and the ability to defer their Social Security benefits.

Limitations of the study stem from not being able to adequately mitigate the existence of "justification bias" although there are mixed empirical findings about the existence and magnitude of such bias. It would also be important to include future leads regarding health for the older workers, in addition to measuring past and current health. This study has opened interesting possibilities for future research. For example, it would be interesting to further investigate transitions in and out from full-time employment to the different exit routes. This could indicate whether improvements in health bring retirees (ones who have exited) back into the labor force full time.

Appendix 1

Latent Health Stock

The HRS has a variety of health measures. These include a subjective general measure of individual's self-reported health and relatively more objective measures of health based such as functional limitations (ADL difficulty), medical diagnosis of chronic illnesses, body mass index and health care utilization which are reported in Table 1. Although self-reported health has been widely used in several studies based on survey data, it may be plagued with problems that lead to bias. As discussed earlier in the paper the problems pertaining to self-reported health are first, self-reported measures of health are based on subjective judgments and there is no reason to believe that these judgments are comparable across individuals. Second, since poor health may represent a legitimate reason for a person of working age to be outside the labor force, respondents who are not working may cite health problems as a way to rationalize behavior (the "justification hypothesis"). The alternative to using self-reported health could be substituting it by relatively more objective measures of health.[26] But these measures may also be self-reported or assessed by the interviewer such that they are not superior indicators of an individual's health (Bound 1991). In order to mitigate the problems associated with self-reported measure of individual health, I have defined a latent health stock variable. Following Bound (1991) and implemented in Bound et al. (1998), a model of self-reported health as a function of relatively more objective measures of health (reported in Table 1) is estimated to create a latent health stock.[27] Then the predicted value for the latent health stock is used as a regressor in hazard analysis.

I adopted the approach of Rice et al. (2010) and used an ordered probit model to estimate self-reported health, where the ordered measure of self-reported health (1 = poor, 2 = fair, 3 = good or very good and 4 = excellent) is regressed

[26] Dwyer and Mitchell (1999), McGarry (2004) and Miah and Wilcox-Gök (2007)

[27] This approach has been used in studies like Disney et al. (2006) that uses objective health indicators as well as other personal characteristics and Jones et al. (2010) which uses only the objective health indicators, in creating the latent health stock.

on 16 relatively more objective physical and mental health explanatory variables and healthcare utilization. The predicted value of the outcome from this estimation is the latent health stock variable which is used as a regressor in proportional hazard model in the main body of the paper. Accordingly, a lower level of health status is given by a smaller value of the latent health stock while a higher level of health status is given by a larger value of the latent health stock. Table 15 presents the marginal effects of the objective health measures for the four different responses (cut points) for self-reported health in an ordered probit model. All objective measures have a statistically significant impact on an individual's self-report of health but each measure weighs differently across the four response categories. In Table 15, column (1) positive marginal effects imply incidence of functional limitations, chronic conditions, depression, higher BMI, more nights spent at hospital, more doctor office visits and higher out of pocket medical expenditure will increase the probability with which an individual is predicted to be in the lowest health category (poor). Similarly, in column (4) the negative marginal effects signify incidence of functional limitations, chronic conditions, depression, higher BMI, more nights spent at hospital, more doctor office visits and higher out of pocket medical expenditure will increase the probability with which an individual is predicted to be in the highest health category (excellent). The same holds true for the marginal effects in the other columns. An increase in latent health stock implies a change from prediction of poor health to better health.

Principal Component Analysis (PCA).

Principal components analysis is a method for detecting a small number of uncorrelated variables, called "principal components", from a large dataset. The objective of principal components analysis is to explain the maximum amount of variance with the minimum number of principal components. PCA analyzes a dataset representing observations described by several variables, which are, in general, inter-correlated. Its goal is to extract the important information from the data and to express this information as a set of new orthogonal variables called *principal components*. The primary goal of principal component analysis is data reduction and addressing multicollinearity. It is a non-parametric technique which has an underlying weakness- data reduction due to PCA leads to loss of information. The association between the components and the original variables is called the component's eigenvalue. In multivariate analysis, the correlation between the component and the original variables is called the component loadings (factor loadings) which are analogous to correlation coefficients, squaring them give the amount of explained variation. Therefore the component loadings tell us how much of the variation in a variable is explained by the component.

In this paper, the main purpose of using principal component analysis is to lend more objectivity to health measures. According to the theory of health production function, an individual's health is a durable good which depends on several factors, some of which may be influenced by an individual. Hence health status of an individual does not solely depend on incidence of physical and mental diseases but on factors like utilization of healthcare inputs, lifestyle behavioral practices, job characteristics, genetic elements etc. Accordingly, I use twenty-eight interrelated variables that are likely to influence the health status of an individual. These variables are reported in Appendix Table 11. In addition to the standard physical and mental health measures (which includes ADL difficulties, other mobility difficulties, chronic illnesses, depression and cognitive problems), I have included information on memory related diseases (Dementia and Alzheimer), healthcare utilization (hospital stay, nursing home stay, doctor office visits and out of pocket medical expenditure), lifestyle factors (smoking, drinking behavior, exercising), job related characteristics (stress, physical effort at work) and genetic information (proxied by average age of parents). From variables, PCA yields 28 factors or principal components. Out of these 28 extracted components only eight with Eigen value greater than 1 are retained (reported in Appendix Table 10). This is known as the "*Kaiser-Gutman*" *Rule*. The sum of all Eigen values is equal to number of included variables. In Table 10, 'Difference' column shows the difference in two consecutive Eigen values. 'Proportion' represents the relative weight of each factor in the total variance. For example Factor 1 (Chronic Condition Factor) explains 13% of the total variance. 'Cumulative Proportion Explained' shows the amount of variance explained by n + (n-1) factors. For example Factor 1 (Chronic Conditions Factor) and Factor 2 (Functional Limitations) explain 22% of total variance. Similarly the eight chosen factors together explain 54% of the total variance. Table 11 shows the pattern matrix which gives a clearer picture of the relevance of each variable in a factor. Factor loadings are the weights and correlations between each variable and the factor. The higher the load the more relevant is the variable in defining the factor's dimensionality. A positive sign indicates a positive relation between the variable and the factor and a negative value indicates an inverse impact on the factor. Uniqueness is the variance that is 'unique' to the variable and not shared with other variables. Each factor is named keeping in mind the variables that load heavily in them. As illustrated in Appendix Table 11, number of chronic conditions load most heavily in Factor 1. Number of ADL difficulties and Mobility Difficulties define Factor 2, but the former has higher loading or correlation with the factor. Hence ADL difficulties are more important that mobility difficulties. Similarly, nights spent at hospital and nursing home define Factor 3, incidence of memory related diseases (Dementia and Alzheimer) and total cognition score defines Factor 4. The heaviest factor loadings for each factor are shaded in grey in Appendix Table 11.

These factors are orthogonal to each other which means they are not correlated to each other. Based on factor loadings I have labeled the factors as Factor 1: Has chronic conditions, Factor 2: Has functional limitation, Factor 3: Hospital stay, Factor 4: Has cognitive functioning problems, Factor 5: Has depression, Factor 6: Lack of physical exercise, Factor 7: Has cancer, and Factor 8: Has lifestyle behavioral problems. Principal components are used because several variables together rather than alone define an interpretable concept. The predicted value of the factors are then used in hazard analysis. Without using PCA it would not be possible to disentangle the causal effect of the health measures since they are highly interrelated. Although uncorrelated factors created through PCA are valuable for empirical model in the paper, there are limitations like loss of information due to aggregation and difficulty in interpretation of regressions coefficients.

Appendix 2

Table 8 Frequency distribution for spells (1992–2010)

Spells	Frequency	Percent	Cumulative
Starting in Wave 1	4,128	26.73	95.21
	Adds 740		
Starting in Wave 2	166	1.07	96.28
Starting in Wave 3	127	0.82	97.11
Starting in Wave 4	92	0.6	97.7
Starting in Wave 5	94	0.61	98.31
Starting in Wave 6	90	0.58	98.89
Starting in Wave 7	70	0.45	99.35
Starting in Wave 8	65	0.42	99.77
Starting in Wave 9	36	0.23	100
Total Spells	4,868		
Total Observations (person-waves)	15,442		
Still Working Full-Time (person-waves)	10,574		

Using data from the Health and Retirement Study (1992–2010) spells of full-time employment are set up for 4128 individuals, at least 50 years of age and working full-time in 1992. Exits are first considered absorbing state- single time exits such that there are 4128 unique spells of full-time employment for 4128 individuals. On considering multiple exits from full-time work, another 740 spells are added later such that there are 4868 spells of full-time employment

Table 9 Kaplan meir baseline hazard rates by different exit routes

Waves	Year	All risks: A l l routes	Competing risks				
			Exit route 1: Complete retirement	Exit route 2: Partial retirement	Exit Route 3: Part-time work	Exit Route 4: unemployment	Exit Route 5: Disability and Not in labor force
1	1992	21.5%	9.6%	5.4%	5.7%	1.6%	1.6%
2	1994	22.8%	10.8%	5.6%	3.5%	1.1%	1.8%
3	1996	23.8%	13.0%	6.6%	2.3%	0.6%	1.1%
4	1998	26.4%	15.1%	7.6%	2.7%	0.5%	0.7%
5	2000	29.6%	17.1%	8.4%	4.6%	0.7%	0.8%
6	2002	33.4%	18.8%	10.4%	3.6%	0.6%	0.1%
7	2004	33.9%	18.5%	10.6%	4.3%	0.1%	0.5%
8	2006	28.9%	13.2%	9.4%	3.8%	2.1%	0.4%
9	2008	32.1%	16.8%	10.2%	2.2%	2.2%	0.7%

The table represents the baseline hazard rates associated with different kinds of exits- all routes (all risks) lumped together versus five different routes of exit (competing risks) over time. The cumulative hazard rate for exit from full-time employment via any route in general as well as individual routes increases over time. This is distinctly observed for complete retirement and partial retirement

Table 10 Principal component analysis

Factor	Eigen Value	Difference	Proportion	Cumulative Proportion Explained
Factor1	3.57	1.03	0.13	0.13
Factor2	2.54	0.12	0.09	0.22
Factor3	2.42	0.20	0.09	0.30
Factor4	1.22	0.05	0.08	0.38
Factor5	1.17	0.05	0.04	0.43
Factor6	1.12	0.09	0.04	0.47
Factor7	1.02	0.01	0.04	0.50
Factor8	1.02	0.03	0.04	0.54
Factor9	0.98	0.01	0.04	0.57
Factor10	0.97	0.03	0.04	0.61
Factor11	0.94	0.01	0.03	0.64
Factor12	0.93	0.07	0.03	0.67
Factor13	0.86	0.03	0.03	0.71
Factor14	0.83	0.01	0.03	0.74
Factor15	0.82	0.02	0.03	0.76
Factor16	0.81	0.04	0.03	0.79
Factor17	0.77	0.01	0.03	0.82
Factor18	0.76	0.01	0.03	0.85
Factor19	0.75	0.05	0.03	0.88
Factor20	0.70	0.01	0.03	0.90
Factor21	0.70	0.09	0.03	0.93
Factor22	0.61	0.09	0.02	0.95
Factor23	0.52	0.20	0.02	0.97
Factor24	0.32	0.02.	0.01	0.98
Factor25	0.30	0.09	0.01	0.99
Factor26	0.21	0.09	0.01	1.00
Factor27	0.12	0.10	0.00	1.00
Factor28	0.01		0.00	1.00

From twenty eight diverse health outcome variables. Eight factors with Eigen value greater than 1 are generated using Principal Component Factor Analysis which is used to create the health indices used as explanatory variable in the hazard model

Proportion indicates the relative weight of each factor in the total variance. For example Factor 1 (Chronic Condition Factor) explains 13% of the total variance

Cumulative Proportion Explained shows the amount of variance explained by n + (n-1) factors. For example Factor 1 (Chronic Conditions Factor) and Factor 2 (Functional Limitations) explain 22% of total variance. Similarly the eight chosen factors together explain 54% of the total variance

Table 11 Principal component analysis- rotation (pattern matrix)

INCLUDED VARIABLES	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Uniqueness
Self-Reported Health	0.25	0.29	0.12	0.33	0.14	-0.15	0.03	0.07	0.29
No. of ADL Difficulties	-0.15	0.76	0.01	0.01	-0.03	-0.01	0.09	-0.04	0.47
No. of Mobility Difficulties	0.05	0.72	0.01	0.00	0.10	-0.20	-0.07	0.01	0.38
Whether Health Limits Work	0.01	0.60	0.18	-0.05	-0.01	0.09	-0.05	0.03	0.57
No. Of Chronic Conditions	0.96	-0.04	0.07	-0.03	0.06	-0.02	0.15	0.00	0.02
Has High BP	0.69	-0.16	-0.02	-0.01	-0.08	-0.35	-0.05	-0.01	0.44
Has Diabetes	0.34	-0.12	0.15	0.20	-0.06	-0.31	0.03	0.01	0.60
Has Heart Disease	0.35	0.05	0.50	-0.11	-0.07	0.20	-0.12	0.06	0.51
Has Lung Disease	0.36	0.05	0.04	-0.03	0.08	0.41	-0.10	0.20	0.60
Had Stroke	0.12	-0.08	0.50	0.01	0.01	0.05	-0.23	-0.12	0.24
Had Cancer	0.18	-0.06	-0.08	-0.05	-0.06	0.05	0.77	0.06	0.39
Has Arthritis	0.62	0.20	-0.30	-0.08	0.06	0.10	0.10	-0.08	0.46
Has Psychological Problem	0.25	-0.12	-0.10	0.04	0.58	0.19	0.17	-0.03	0.51
Depression	-0.08	0.20	-0.06	0.32	0.62	-0.06	-0.05	-0.03	0.46
Total Cognition Score	0.06	0.02	0.01	-0.77	-0.05	0.03	0.06	-0.10	0.43
Alzheimer's	0.13	0.02	0.01	0.69	0.18	0.02	0.06	0.11	0.33
Dementia	0.11	0.01	0.01	0.71	0.12	0.02	0.02	0.06	0.31
BMI	0.28	0.25	-0.15	0.02	-0.21	0.54	-0.07	-0.12	0.46
Out of Pocket Expenditure	-0.02	-0.14	0.35	0.00	0.23	-0.13	0.27	-0.04	0.65
No. of Nights Hospital Stay	0.07	0.13	0.68	-0.01	-0.09	0.07	0.10	0.02	0.51
No. of Doc Visits	0.07	0.25	0.26	-0.10	-0.01	0.01	0.53	0.06	0.49
No. of Nights Nursing Home Stay	0.05	0.10	0.67	-0.02	-.0.05	0.04	0.19	0.01	0.38
Ever Smoked	0.02	-0.03	0.04	0.25	-0.04	0.07	0.01	0.79	0.35
Drinks Alcohol	-0.07	0.02	-0.09	-0.26	-0.06	-0.12	0.06	0.65	0.46
Does Vigorous Physical Activity	0.01	-0.09	0.07	0.15	-0.20	-0.56	0.07	-0.19	0.54
Has Stress At Work	-0.01	0.03	-0.01	-0.31	0.57	-0.05	-0.23	-0.07	0.40
Job Requires Physical Effort	-0.02	0.03	-0.13	0.50	0.01	0.25	-0.07	-0.09	0.57
Average Age of Parents	-0.01	-0.02	-0.01	0.01	0.02	0.01	-0.05	0.01	0.11

Information on chronic conditions and specific mental problems are based on medical diagnosis in previous two years and intake of prescription drug for the same. The shaded cells show the health variables that load heavily in the respective factors. The sign of the respective variables in each factor indicates how they weigh in that factor. A positive sign indicates a positive relation between the variable and the factor while a negative sign indicates an inverse relationship. Each factor is named keeping in mind the variables that load heavily in them. These factors are orthogonal to each other which means they are not correlated to each other

Table 12 Competing risk model (weibull distribution) for initial exit from full-time employment (with frailty)

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
Latent Health Stock: Self-Reported Health	0.812***	13.34%	0.785***	6.62%	1.106	4.26%	0.971	2.46%	0.971	0.72%	0.425***	0.36%
	(0.042)		(0.047)		(0.120)		(0.147)		(0.244)		(0.078	
Factor 1: Chronic Conditions	0.960	15.78%	1.004	8.48%	0.945	3.65%	0.892	2.26%	0.793	0.59%	0.939	0.80%
	(0.029)		(0.035)		(0.056)		(0.079)		(0.123)		(0.084)	
Factor 2: Functional Limitation	1.126***	18.51%	1.146***	9.67%	1.084	4.18%	0.916	2.32%	1.225	0.91%	1.278***	1.08%
	(0.026)		(0.032)		(0.057)		(0.074)		(0.158)		(0.078)	
Factor 3: Hospital Stay	1.016	16.70%	1.010	8.52%	0.931	3.59%	1.099	2.79%	1.061	0.79%	1.135	0.97%
	(0.027)		(0.032)		(0.054)		(0.082)		(0.145)		(0.092)	
Factor 4: Cognitive Functioning	1.089**	17.90%	1.092	9.23%	1.003	3.87%	1.146	2.91%	0.770	0.57%	1.338***	1.13%
	(0.038)		(0.047)		(0.069)		(0.116)		(0.135)		(0.151)	
Factor 5: Depression	1.112***	18.28%	1.096***	9.25%	0.916	3.54%	1.207**	3.07%	1.262*	0.93%	1.329	1.13%
	(0.029)		(0.036)		(0.054)		(0.093)		(0.161)		(0.094)	
Factor 6: Physical Exercise	0.963	15.83%	0.963	8.12%	1.069	4.12%	0.999	2.54%	0.714**	0.53%	1.035	0.88%
	(0.026)		(0.031)		(0.056)		(0.078)		(0.099)		(0.087)	
Factor 7: Cancer	0.915***	15.04%	0.915***	7.72%	0.929	3.59%	0.845**	2.14%	0.775*	0.57%	0.919	0.79%
	(0.025)		(0.030)		(0.049)		(0.070)		(0.119)		(0.081)	
Factor 8: Life Style	1.142***	18.78%	1.154***	9.74%	1.172***	4.52%	1.068	2.71%	0.995	0.74%	1.059	0.90%

Table 12 (continued)

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
	(0.034)		(0.040)		(0.066)		(0.089)		(0.138)		(0.098)	
# Spells	4128		4128		4128		4128		4128		4128	
# Exits	3272		1680		769		505		148		170	
Avg. Spell Length	4.82		4.82		4.82		4.82		4.82		4.82	
Avg. Exit Probability	16.44%		8.44%		3.86%		2.54%		0.74%		0.85%	

Table 12 uses data from the Health and Retirement Study (1992–2010) for a proportional hazard model in competing risks framework for individuals who were at least 50 years old and worked full-time in 1992. The baseline hazard is estimate using the Weibull distribution where the shape parameter p is approximately close to 2. It reports hazard ratios and absolute effect for the initial exits scenario. Absolute Effect = Hazard Ratio * Average Exit Probability via that route, where average exit probability = no. of exits/(no. of spells * average spell length). The dependent variable is duration of full-time employment while the independent variables are health outcome indices constructed by principal component factor analysis. The competing risks in this framework are the different routes of exit from full-time employment. The combined risk model has also been reported in panel A. Controlling for the socio-demographic variables reported in Table 2. Also controlling occupations, the wave dummies, census regions, expected mortality and work and health status of spouse, which has not been reported here. Further, it controls for frailty (unobserved heterogeneity). The standard errors are reported in parentheses and *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Likelihood ratio test rejects the null hypothesis of No Frailty such that there is statistically significant level of unobserved heterogeneity in the data. The shared frailty is estimated assuming an underlying gamma distribution

Table 13 Competing risk model (cox proportional hazard) for multiple exit from full-time employment

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
Latent health stock: Self-reported health	0.843***	14.28%	0.778***	6.49%	1.080	4.64%	0.970	2.66%	0.933	0.65%	0.511***	0.43%
	(0.031)		(0.040)		(0.089)		(0.089)		(0.164)		(0.064)	
Factor 1: Chronic conditions	1.038*	17.58%	1.065**	8.89%	1.051	4.51%	1.018	2.79%	0.889	0.62%	1.001	0.85%
	(0.022)		(0.033)		(0.047)		(0.054)		(0.091)		(0.083)	
Factor 2: Functional limitation	1.065***	18.04%	1.111***	9.28%	1.024	4.40%	0.898**	2.46%	1.112	0.77%	1.200***	1.02%
	(0.019)		(0.027)		(0.044)		(0.049)		(0.095)		(0.069)	
Factor 3: Hospital stay	1.016	17.21%	1.016	8.48%	0.933	4.01%	1.089*	2.98%	1.024	0.71%	1.159	0.98%
	(0.020)		(0.028)		(0.043)		(0.050)		(0.100)		(0.088)	
Factor 4: Cognitive functioning	1.039	17.60%	1.053	8.79%	0.943	4.05%	1.026	2.81%	0.892	0.62%	1.307**	1.11%
	(0.027)		(0.040)		(0.051)		(0.063)		(0.102)		(0.139)	
Factor 5: Depression	1.059***	17.93%	1.064**	8.88%	0.931	4.00%	1.028	2.81%	1.166*	0.81%	1.256***	1.07%
	(0.021)		(0.031)		(0.042)		(0.047)		(0.095)		(0.079)	
Factor 6: Physical exercise	1.019	17.26%	0.970	8.10%	1.094**	4.69%	1.073	2.94%	0.889	0.62%	1.123	0.96%
	(0.020)		(0.028)		(0.045)		(0.053)		(0.084)		(0.091)	
Factor 7: Cancer	0.984	16.66%	0.964	8.04%	0.983	4.22%	1.009	2.76%	0.958	0.67%	0.981	0.84%
	(0.018)		(0.026)		(0.038)		(0.046)		(0.095)		(0.073)	
Factor 8: Life style	1.041*	17.64%	1.083***	9.04%	1.053	4.52%	0.989	2.71%	1.020	0.71%	0.959	0.82%

Table 13 (continued)

Health outcomes	All risks		Competing risks									
			Exit route 1: Complete retirement		Exit route 2: Partial retirement		Exit route 3: Part-time work		Exit route 4: Unemployment		Exit route 5: Disability and Not in labor force	
	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect	Hazard ratio	Absolute effect
	(0.021)		(0.033)		(0.044)		(0.048)		(0.096)		(0.082)	
# Spells	4868		4868		4868		4868		4868		4868	
# Exits	3817		1883		968		617		157		192	
Avg. spell length	4.63		4.63		4.63		4.63		4.63		4.63	
Avg. exit probability	16.94%		8.35%		4.29%		2.74%		0.70%		0.85%	

Table 13 uses data from the Health and Retirement Study (1992–2010) for a proportional hazard model in competing risks framework for individuals who were at least 50 years old and worked full-time in 1992. The underlying baseline hazard is assumed not to have any parametric functional form It reports hazard ratios and absolute effect for the multiple exits scenario. Absolute Effect = Hazard Ratio * Average Exit Probability via that route, where average exit probability = no. of exits/(no. of spells * average spell length). The dependent variable is duration of full-time employment while the independent variables are health outcome indices constructed by principal component factor analysis. The competing risks in this framework are the different routes of exit from full-time employment. The combined risk model has also been reported in panel A. Controlling for the socio-demographic variables reported in Table 2. Also controlling occupations, the wave dummies, census regions, expected mortality and work and health status of spouse, which has not been reported here. Further, it controls for frailty (unobserved heterogeneity). The standard errors are reported in parentheses and

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.Likelihood ratio test rejects the null hypothesis of No Frailty such that there is statistically significant level of unobserved heterogeneity in the data. The shared frailty is estimated assuming an underlying gamma distribution

Table 14 Competing risk model for exiting fulltime employment due to health changes allowing for multiple spells of employment (with frailty)

Health change	(A) All risks		Competing risks									
			Exit Route 1: Complete Retirement		Exit Route 2: Partial Retirement		Exit Route 3: Part-time Work		Exit Route 4: Unemployment		Exit Route 5: Disability and Not in Labor force	
	Hazard Ratio	Absolute Effect	Hazard Ratio	Absolute Effect	Hazard Ratio	Absolute Effect	Hazard Ratio	Absolute Effect	Hazard Ratio	Absolute Effect	Hazard Ratio	Absolute Effect
Change in Overall Self-Reported Health	1.231***	20.85%	1.212***	10.12%	1.114	4.78%	1.125	2.86%	1.187	0.87%	1.099***	0.93%
	(0.041)		(0.060)		(0.080)		(0.102)		(0.189)		(0.222)	
Change in Count of Functional Limitations	1.095	18.54%	1.169	9.76%	1.146	5.31%	1.099	2.80%	1.371	1.01%	1.631**	1.39%
	(0.103)		(0.146)		(0.282)		(0.273)		(0.425)		(0.291)	
Change in Count of Chronic Conditions	1.109**	18.79%	1.085**	9.06%	1.029	4.41%	1.194	3.03%	1.250	0.93%	1.354*	1.15%
	(0.059)		(0.078)		(0.105)		(0.121)		(0.234)		(0.201)	
Onset of Memory Related Disease	1.101*	18.65%	1.127*	9.41%	1.067	4.57%	0.585	1.49%	1.384	1.02%	1.761*	1.50%
	(0.112)		(0.153)		(0.257)		(0.311)		(0.517)		(0.406)	
# Spells	4868		4868		4868		4868		4868		4868	
# Exits	3817		1883		968		505		148		170	
Avg. spell length (in 2 year waves)	4.63		4.63		4.63		4.82		4.82		4.82	
Avg. exit probability	16.94%		8.35%		4.29%		2.54%		0.74%		0.85%	

Health changes between waves are accounted for self-reported overall health (individuals who report health being same as previous wave to either much better or much worse in next wave; increase in count of ADL difficulties (functional limitations) and chronic conditions from 0 in previous wave to 5 or reverse and onset of memory related diseases between waves. Controlling for the socio- demographic variables reported in Table 2. Also controlling occupations, the wave dummies, census regions, expected mortality and work and health status of spouse, which has not been reported here. Unobserved heterogeneity has also been controlled. The standard errors are reported in parentheses and *** $p < 0.01$, ** $p < 0.05$, * $p < 0$

Table 15 Ordered probit regression of self-reported health (marginal effects)

Variables	Self-reported health = 1	Self-reported health = 2	Self-reported health = 3	Self-reported health = 4
Has ADL difficulty	0.0160***	0.00507***	−0.0166***	−0.00838***
	(12.93)	(12.79)	(−12.99)	(−12.92)
Has multiple chronic conditions	0.0154***	0.00486***	−0.0159***	−0.00804***
	(5.986)	(5.961)	(−5.989)	(−5.985)
Has high BP	0.0337***	0.0107***	−0.0347***	−0.0177***
	(10.95)	(10.63)	(−10.94)	(−10.84)
Has diabetes	0.0662***	0.0115***	−0.0665***	−0.0293***
	(18.60)	(29.68)	(−19.10)	(−21.67)
Has lung disease	0.0656***	0.00925***	−0.0656***	−0.0279***
	(15.42)	(24.89)	(−15.96)	(−18.84)
Has heart disease	0.0638***	0.0130***	−0.0645***	−0.0293***
	(18.88)	(27.16)	(−19.27)	(−21.18)
Had stroke	0.00871**	0.00253**	−0.00896**	−0.00441**
	(2.159)	(2.360)	(−2.168)	(−2.225)
Has arthritis	0.0169***	0.00545***	−0.0175***	−0.00891***
	(5.476)	(5.349)	(−5.468)	(−5.430)
Has cancer	0.0290***	0.00711***	−0.0296***	−0.0139***
	(8.178)	(11.02)	(−8.277)	(−8.887)
BMI	0.000860***	0.000272***	−0.000887***	−0.000449***
	(5.752)	(5.715)	(−5.748)	(−5.749)
Has depression	0.0334***	0.0105***	−0.0344***	−0.0174***
	(68.22)	(44.66)	(−69.04)	(−65.24)
Total cognitive index score	−0.00683***	−0.00216***	0.00705***	0.00357***
	(−42.25)	(−33.83)	(42.07)	(41.36)
Has psychological problem	0.00198***	0.00645***	−0.0981***	−0.1025***
	(5.56)	(5.66)	(−4.87)	(−5.92)
No. of nights at hospital	0.0306***	0.0117***	−0.0319***	−0.0173***
	(16.28)	(13.53)	(−16.12)	(−14.98)
No. of Doc visits	0.0183***	0.00597***	−0.0189***	−0.00968***
	(11.65)	(11.10)	(−11.63)	(−11.47)
Out of pocket medical expenditure	0.0376***	0.0119***	−0.0388***	−0.0197***
	(23.39)	(21.68)	(−23.40)	(−23.13)
No. of observations	15,442	15,442	15,442	15,442

T statistics are reported in parentheses and * $p < 0.1$, ** $p < 0.05$, *** $p < 0.0$

Table 16 Descriptive statistics for some health outcome variables measured in 1992

Variables	Individuals working fulltime in 1992 who subsequently exit for the first time via:						Difference in sample means				
	(1) Working full-time	(2) Complete retirement route	(3) Partial retirement route	(4) Part-time work route	(5) Unemployment route	(6) Disability/Not in labor force route	Difference in means (1)–(2)	Difference in means (1)–(3)	Difference in means (1)–(4)	Difference in means (1)–(5)	Difference in means (1)–(6)
Health outcomes											
Has poor health	0.0134	0.0430	0.0276	0.0335	0.0143	0.0896	−0.0296***	0.0142***	−0.0201*	−0.0009	−0.0762***
Has good health	0.6170	0.5072	0.5743	0.5789	0.6571	0.3582	0.110***	0.0427	0.0381	−0.0401	0.259***
Has ADL difficulty	0.0061	0.0258	0.0068	0.0144	0.0143	0.0299	−0.0197***	−0.0007	−0.0083	−0.0082	−0.0238*
Has multiple chronic conditions	0.5607	0.6619	0.6081	0.5455	0.0068	0.0144	−0.101***	−0.0474	0.0153	0.0322	−0.0363
Has high BP	0.2992	0.3496	0.2432	0.3110	0.6081	0.5455	−0.0503	0.0560	−0.0118	−0.0293	−0.0142
Has diabetes	0.0603	0.1117	0.0878	0.0766	0.0286	0.1194	−0.0515***	−0.0276	−0.0163	0.0317	−0.0591*
Has cancer	0.0384	0.0401	0.0473	0.0574	0.0143	0.0299	−0.0018	−0.0089	−0.0191	0.0241	0.0085
Has heart disease	0.0721	0.1175	0.1216	0.0813	0.0714	0.0896	−0.0453**	−0.0495*	−0.0092	0.0007	−0.0174
Has lung disease	0.0244	0.0430	0.0946	0.0335	0.0322	0.0299	−0.0186*	−0.0702***	−0.0091	−0.0078	−0.0055
Has arthritis	0.2886	0.3467	0.2770	0.2727	0.2429	0.4030	−0.0581*	0.0116	0.0159	0.0457	−0.114*
Has depression	0.4021	0.4298	0.3716	0.5024	0.5000	0.7463	−0.0277	0.0305	−0.100**	−0.0979	−0.344***
Total cognition index score	24.6387	23.5931	23.4797	23.6890	24.3143	24.0746	1.046***	1.159**	0.950**	0.3240	0.5640
Has psychological problem	0.0292	0.0315	0.0135	0.0526	0.0286	0.0746	−0.0023	0.0157	−0.0234	0.0007	−0.0454*
No. of nights at hospital*	0.3881	0.7307	0.5338	0.3828	0.2286	0.5970	−0.343**	−0.1457	0.0054	0.1600	−0.2090

Table 16 (continued)

Variables	Individuals working fulltime in 1992 who subsequently exit for the first time via:						Difference in sample means				
	(1) Working full-time	(2) Complete retirement route	(3) Partial retirement route	(4) Part-time work route	(5) Unemployment route	(6) Disability/Not in labor force route	Difference in means (1)–(2)	Difference in means (1)–(3)	Difference in means (1)–(4)	Difference in means (1)–(5)	Difference in means (1)–(6)
No. of doctor visits*	3.1717	4.4241	3.5743	2.6459	2.8286	4.5821	−1.252***	−0.4026	0.5260	0.3430	−1.410*
Out of pocket medical expenditure	0.0332	0.0516	0.0431	0.0526	0.0429	0.1045	−0.0184	-0.0099*	−0.0195	−0.0097	−0.0713**
No. of observations	3285	349	148	209	70	67					

Here * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ indicate the statistical significance of the health measures for the test of equality of means for each exit route with full-time employment

Table 17 Descriptive statistics for socio-demographic variables measured in 1992

Variables	Individuals working fulltime in 1992 who subsequently exit for the first time via:						Difference in sample means				
	(1) Working full-time	(2) Complete retirement route	(3) Partial retirement route	(4) Part-time work route	(5) Unemployment route	(6) Disability/ Not in labor force route	Difference in means (1)–(2)	Difference in means (1)–(3)	Difference in means (1)–(4)	Difference in means (1)–(5)	Difference in means (1)–(6)
Socio-demographic											
Age	55.6658	58.6161	59.4932	55.9330	55.3143	54.8209	−2.950***	−3.827***	−0.2670	0.3510	0.845*
Female	0.4280	0.3639	0.2838	0.6555	0.4429	0.7164	0.0641*	0.144***	−0.227***	−0.0149	−0.288***
Black	0.1370	0.1461	0.0946	0.2105	0.2000	0.2239	−0.0092	0.0424	−0.0735**	−0.0630	−0.0869*
Hispanic	0.0192	0.0057	0.0270	0.0239	0.0143	0.0123	0.0134	−0.0079	−0.0048	0.0049	0.0375
Other race	0.0621	0.0458	0.0473	0.1053	0.1143	0.1642	0.0163	0.0148	−0.0432*	−0.0522	−0.102***
Married	0.7473	0.7736	0.8378	0.7225	0.6429	0.5821	−0.0263	−0.0905*	0.0248	0.104*	0.165**
Years of schooling	12.8429	12.3209	12.8041	11.9282	12.7000	11.4328	0.522**	0.0389	0.915***	0.1430	1.410***
Blue collar worker	0.2843	0.3295	0.2770	0.2344	0.3571	0.3881	−0.0452	0.0073	0.0499	−0.0728	−0.1040
Has religious preference	0.9470	0.9599	0.9662	0.9569	0.9143	0.9254	−0.0129	−0.0192	−0.0099	0.0327	0.0217
Native	0.9084	0.9599	0.9595	0.8708	0.8571	0.8657	−0.0515**	−0.0511*	0.0376	0.0512	0.0427
Mother's age	73.5627	73.4556	73.4082	72.9563	70.3333	70.5373	0.1070	0.1550	0.6060	3.229*	3.0250
Father's age	70.6056	70.0000	69.7619	69.0588	68.0882	69.9105	0.6060	0.8440	1.5470	2.5170	0.6950
Father's education (>8 yrs)	0.5656	0.5444	0.5541	0.5502	0.7429	0.5075	0.0212	0.0115	0.0154	−0.177**	0.0581
Mother's education (>8 yrs)	0.6134	0.6046	0.6284	0.5837	0.7143	0.5522	0.0088	−0.0150	0.0297	−0.1010	0.0612
Individual income	31,605.7300	28,862.5400	31,604.6600	19,488.2700	27,437.9100	18,031.2500	2743.2000	1.0640	12,117.5***	4167.8000	13,574.5***
Household income	54,537.2400	51,095.8200	57,483.9500	51,897.8200	46,381.0000	35,770.4800	3441.4000	−2946.7000	2639.4000	8156.2000	18,766.8***
Has govt. insurance	0.0539	0.1060	0.1081	0.0335	0.0286	0.1045	−0.0521***	−0.0542**	0.0204	0.0253	−0.0506
Has pvt. insurance	0.1382	0.1347	0.1959	0.2105	0.0857	0.1194	0.0035	−0.0577*	−0.0723**	0.0525	0.0188

Table 17 (continued)

Variables	Individuals working fulltime in 1992 who subsequently exit for the first time via:						Difference in sample means				
	(1) Working full-time	(2) Complete retirement route	(3) Partial retirement route	(4) Part-time work route	(5) Unemployment route	(6) Disability/ Not in labor force route	Difference in means (1)–(2)	Difference in means (1)–(3)	Difference in means (1)–(4)	Difference in means (1)–(5)	Difference in means (1)–(6)
Has employer insurance	0.8667	0.8596	0.7703	0.6555	0.7571	0.7015	0.0071	0.0964***	0.211***	0.110**	0.165***
Has Non-employer insurance	0.0606	0.0802	0.1149	0.1531	0.0714	0.1343	−0.0197	−0.0543**	−0.0925***	−0.0109	−0.0737*
No insurance	0.0728	0.0602	0.1149	0.1914	0.1714	0.1642	0.0126	−0.0421	−0.119***	−0.0987**	−0.0914**
Has pension DC	0.2018	0.1576	0.1622	0.1196	0.2286	0.1194	0.0442*	0.0397	0.0822**	−0.0267	0.0824
Has pension DB	0.2734	0.3410	0.2230	0.1388	0.1571	0.2687	−0.0676**	0.0504	0.135***	0.116*	0.0047
Has pension both	0.1787	0.1920	0.1554	0.0574	0.0714	0.0746	−0.0133	0.0233	0.121***	0.107*	0.104*
Don't know about pension	0.0079	0.0057	0.0068	0.0096	0.0143	0.0149	0.0022	0.0012	−0.0017	−0.0064	−0.0070
No. of observations	3285	349	148	209	70	67					

Here * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ indicate the statistical significance of the health measures for the test of equality of means for each exit route with full-time employment

References

Anderson KH, Burkhauser RV (1985) The retirement-health nexus: a new measure of an old puzzle. J Hum Resour 20(3):315–330. doi:10.2307/145884

Bazzoli GJ (1985) The early retirement decision: new empirical evidence on the influence of health. J Hum Resour 20(2):214–234. doi:10.2307/146009

Benítez-Silva H, Buchinsky M, Man Chan H, Cheidvasser S, Rust J (2004) How large is the bias in self-reported disability? J Appl Econ 19(6):649–670

Borgan Ø, Langholz B (2007) Using martingale residuals to assess goodness-of-fit for sampled risk set data. In: Nair V (ed) Advances in statistical modeling and inference. essays in honor of Kjell A Doksum. World Scientific Publishing, Singapore, p 65–90

Boskin MJ (1977) Social security and retirement decisions. Econ Inq 15(1):1–25

Boskin MJ, Kotlikoff LJ, Puffert DJ, Shoven JB (1987) Social security: a financial appraisal across and within generations: National Bureau of Economic Research Cambridge

Bound J (1991) Self-reported versus objective measures of health in retirement models. J Hum Resour 26(1):106–138. doi:10.2307/145718

Bound J, Schoenbaum M, Stinebrickner TR, Waidmann T (1998) The dynamic effects of health on the labor force transitions of older workers. National Bureau of Economic Research Working Paper Series, No. 6777

Butler JS, Anderson KH, Burkhauser RV (1989) Work and health after retirement: a competing risks model with semiparametric unobserved heterogeneity. Rev Econ Stat 71(1):46–53

Canals-Cerdá J, Gurmu S (2007) Semiparametric competing risks analysis. Econ J 10(2):193–215

Christensen BJ, Kallestrup-Lamb M (2012) The impact of health changes on labor supply: evidence from merged data on individual objective medical diagnosis codes and early retirement behavior. Health Econ 21(S1):56–100

Coile C, Diamond P, Gruber J, Jousten A (2002) Delays in claiming social security benefits. J Public Econ 84(3):357–385

Disney R, Emmerson C, Wakefield M (2006) Ill health and retirement in Britain: a panel data-based analysis. J Health Econ 25(4):621–649. doi:10.1016/j.jhealeco.2005.05.004

Dwyer DS, Mitchell OS (1999) Health problems as determinants of retirement: are self-rated measures endogenous? J Health Econ 18(2):173–193. doi:10.1016/S0167-6296(98)00034-4

Gruber J, Wise DA (2005) Social security programs and retirement around the world: micro-estimation. ILRReview 58(2):86

Irwin M, Artin KH, Oxman MN (1999) Screening for depression in the older adult: criterion validity of the 10-item center for epidemiological studies depression scale (CES-D). Arch Intern Med 159(15):1701–1704

Jones AM, Rice N, Roberts J (2010) Sick of work or too sick to work? Evidence on self-reported health shocks and early retirement from the BHPS. Econ Model 27(4):866–880

Kerkhofs M, Lindeboom M (1995) Subjective health measures and state dependent reporting errors. Health Econ 4(3):221–235

Lindeboom M, Kerkhofs M (2002) Health and work of the elderly: subjective health measures, Reporting Errors and the Endogenous Relationship between Health and Work: Institute for the Study of Labor (IZA)

Marmot M, Banks J, Blundell R, Lessof C, Nazroo J (2002) Health, wealth and lifestyles of the older population in England. Institute for Fiscal Studies, London

Marton J, Ketsche PG, Zhou M (2010) SCHIP premiums, enrollment, and expenditures: a two state, competing risk analysis. Health Econ 19(7):772–791

McGarry K (2004) Health and retirement do changes in health affect retirement expectations? J Hum Resour 39(3):624–648

Meghir C, Whitehouse E (1997) Labour market transitions and retirement of men in the UK. J Econ *79*(2):327–354. doi:10.1016/S0304-4076(97)00026-2

Miah MS, Wilcox-Gök V (2007) Do the sick retire early? Chronic illness, asset accumulation and early retirement. Appl Econ 39(15):1921–1936

Parsons DO (1982) The male labour force participation decision: health, reported health, and economic incentives. Economica 49(193):81–91. doi:10.2307/2553527

Quinn JF (1981) The extent and correlates of partial retirement. The Gerontologist 21(6):634–643

Riphahn RT (1999) Income and employment effects of health shocks. A test case for the German welfare state. J Popul Econ 12(3):363–389

Siddiqui S (1997) The impact of health on retirement behaviour: empirical evidence from West Germany. Health Econ 6(4):425–438. doi:10.1002/(sici)1099-1050(199707)6:4<425::aid-hec284>3.0.co;2-t

Stern S (1989) Measuring the effect of disability on labor force participation. J Hum Resour 24(3):361–395