The Health Consequences of Retirement

Michael Insler

ABSTRACT

This paper examines the impact of retirement on individuals' health. Declines in health commonly compel workers to retire, so the challenge is to disentangle the simultaneous causal effects. The estimation strategy employs an instrumental variables specification. The instrument is based on workers' self-reported probabilities of working past ages 62 and 65, taken from the first period in which they are observed. Results indicate that the retirement effect on health is beneficial and significant. Investigation into behavioral data, such as smoking and exercise, suggests that retirement may affect health through such channels. With additional leisure time, many retirees practice healthier habits.

I. Introduction

How does the decision to retire impact one's health? This question is extremely relevant to the ongoing national policy discussion about major government programs such as Social Security and Medicare. Several developments, including the maturation of the "Baby Boom" generation, the decline in fertility, increased life expectancy, and the higher proportion of deaths due to degenerative ailments, have called into question the continued fiscal viability of these bedrock social programs. Although much of the conversation focuses on their projected costs, there is less emphasis on how changes to the structure of Social Security and Medicare may influence the health of the aging population and how such developments may, in turn, "feed back" as second-order effects on the programs' finances. For instance, an increase in the minimum age to receive Social Security benefits would incentivize workers to retire later, which may consequently alter future health patterns, mortality rates, and Medicare us-

Michael Insler is an assistant professor of economics at the United States Naval Academy. *He is very grateful to Ronni Pavan for extraordinary help and guidance, as well as Mark Bils, Gregorio Caetano, Joshua Kinsler, Nese Yildiz, and the students and faculty at the University of Rochester Applied/Macro lunch seminar for their useful suggestions. He also thanks two anonymous referees for many helpful remarks. Researchers may acquire the data used in this article from RAND's Center for the Study of Aging (http://www.rand.org/labor/aging/dataprod/hrs-data.html). For questions regarding the data, please contact Michael Insler, insler@usna.edu.*
ISSN 0022-166X E-ISSN 1548-8004

age. This paper estimates the *causal* impact of retirement on health. A more complete understanding of the health consequences of retirement will provide a better notion of the economic impact of potential changes to Social Security and Medicare. In general, this new information will allow economists to forecast payout and tax streams more accurately and to model healthcare needs, insurance plans, and labor market transitions with more precision.

A. Background

There is a well-established connection between health and work. Numerous studies have researched the relationship between health and labor supply, a topic that is closely associated with the health insurance market. For many workers, the role of private health insurance can be an important part of their labor market decisions. For instance, there is strong evidence that health insurance is a major factor in the labor force participation of secondary wage earners: Several studies have estimated that spouse-covered health insurance reduces own-labor force participation by seven to 11 percentage points.[1] Most relevant to the current study, however, is the substantial evidence that health insurance is an integral component of workers' retirement decisions. Gustman and Steinmeier (1994) found that employer provision of retiree health insurance delays retirement until the eligibility age for such coverage and accelerates it thereafter. Rust and Phelan (1997) determined that, due to the Medicare eligibility age of 65, men with employer-provided health insurance but without employer-provided health coverage in retirement are less likely to retire before age 65 than those with health insurance that spans retirement. Moreover, if it is true that the strong link between expansive public policies and retirement behavior stems not only from health insurance but also from health *status*, then the latter connection also merits close scrutiny.

Many studies have examined the "retirement-health nexus," particularly the role of health status in individuals' retirement decisions. Anderson and Burkhauser (1985) posed the question of whether retirement plans are driven by economic variables as much as by health. Although their results differed depending on the choice of health measure (self-reported health or mortality), they found that self-assessed health effects on retirement were larger than wage effects. Bazzoli (1985) compared preretirement and postretirement health information, concluding (oppositely) that economic factors, rather than health, had the larger influence on retirement decisions.[2] Dwyer and Mitchell (1999) utilized more detailed, longitudinal health data to argue that health indeed plays a major role in retirement. McGarry (2004) developed a model to examine the effect of health on retirement expectations, concluding that health has a much larger impact than economic factors on the probability of working. This literature reveals that the health-retirement simultaneity issue is pervasive, and an important converse question remains on how retirement may influence health.[3]

What are the implications of retirement's effect on health, should it exist? As stated earlier, there may be second-order effects. Rust and Phelan (1997) indicated that

1. Olson (1998); Buchmueller and Valletta (1999).
2. Bazzoli's results, like those of Anderson and Burkhauser (1985) stemming from mortality, may have been impacted by poorly measured health variables in the data.
3. See Section II for a review of these related studies.

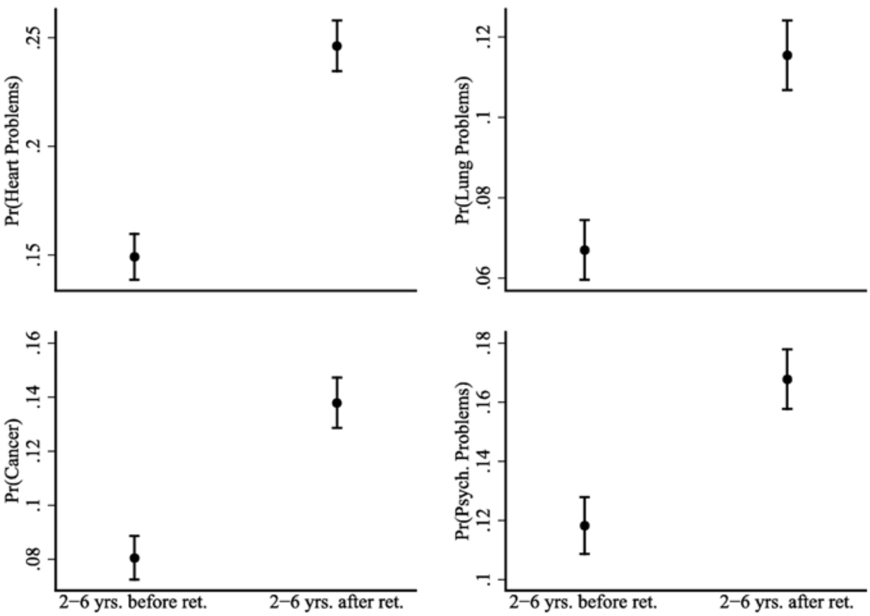

Figure 1
Health Conditions Pre- and Post-Retirement

Notes: The figures plot sample proportions of health conditions and 95 percent confidence intervals, conditional on duration preceding or following retirement. The four ailments presented are heart problems, cancer, lung problems, and psychological disorders.

changes to Social Security and Medicare would influence retirement behavior, but it remains unknown whether subsequent retirement-driven health changes (if they exist) could in turn affect the finances of those programs. Is there any evidence, a priori, that such second-order effects may be significant and hence worth investigating? In 2002, individuals age 65 and over comprised 13 percent of the U.S. population, but they consumed 36 percent of total U.S. personal healthcare expenses (Stanton 2006). Thus it is clear that factors affecting late-career workers' and retirees' health are of great fiscal importance, particularly with respect to discussions regarding the long- term viability of Medicare. Moreover, the five most expensive health conditions in 2002 were heart disease, cancer, trauma, psychological disorders, and pulmonary conditions (Olin and Rhoades 2006). Four of those five conditions (all but trauma) are captured by the health metrics used throughout this paper, and 2002 is near the midpoint of the longitudinal data employed by the main empirical work to come. Figure 1 presents sample proportions and 95 percent confidence intervals for those four conditions (data is from the main sample, which will be described in Section III). Most notably, there is a large and statistically significant increase in the incidence of each ailment, transitioning from late-career workers to recent retirees, jumping by as much as ten percentage points in the case of heart problems. These simple correlations illustrate a potential pitfall of *not* carefully examining the health consequences of retirement. Without a thorough investigation, basic analysis might lead one to believe that retirement exacerbates the most costly and pervasive health conditions, whereas the main results of Section V will show that, in fact, the opposite is true. Such fallacious reasoning could lead to uninformed policy recommendations or poor budgetary projections, potentially harmful to important aspects of fiscal policy, including Medicare, Social Security, and other programs.

B. Framework

In examining this question, the challenge is to properly treat the simultaneous effects that may cloud the true impact of retirement on health. In particular, it is common for workers to retire when they become ill or injured, and as a result, poor health may bring about retirement.[4] There is no strong consensus regarding the converse effect, which may operate in different directions.[5] On the one hand, retirement may lead to a negative lifestyle shock, a loss of ambition, or a general decrease in activity level, expediting the decline in health that naturally accompanies aging. On the other hand, retirement provides retirees with more leisure time so they may address their "health upkeep" needs and experience less job-related stress and strain. This paper constructs an econometric model that allows estimation of retirement's net effect on individuals' health.

Simple empirical analysis of data from the Health and Retirement Study (HRS) suggests that simultaneous effects are indeed present.[6] In this paper, it is crucial to properly measure and interpret the notion of "general health." Section III provides details on the health metrics, but for now, let health be measured by a univariate scale from zero to one. Figure 2 contains sample averages for health levels and health changes, grouped by age. The figure's first panel confirms that, on average, health declines with age, and it highlights the disparity between workers and retirees. Unretired individuals tend to be healthier, while retirees' average health progression exhibits a hump-shaped profile. Young retirees (ages 50–60) generally have exceptionally poor health, suggesting that their early exits from the labor force may be due to severe illnesses or injuries (for instance, it may take a remarkably bad ailment to compel a younger individual to retire early). This phenomenon underscores the simultaneity issue. One might speculate that health changes (as opposed to health levels)—particularly those that occurred after retirement—would escape the simultaneity issue. The second panel of Figure 2 plots average health changes by age, showing that health decay slowly accelerates with aging. The effect is faster for workers, while retirees tend to experience a more stable decline, particularly at higher ages. Analysis of health changes permits identification of retirement's effect on retirees' health evolution *after their retirement*. However, solely taking differences is not enough. For instance, a survey respondent may have suffered a stroke between periods t and $t-1$, concurrently forcing him or her into retirement. The simultaneity problem may be even more

4. Anderson and Burkhauser (1985); McGarry (2004); Dwyer and Mitchell (1999).
5. See Section II on related literature.
6. The RAND HRS Data file is an easy to use longitudinal data set based on the HRS data. It was developed at RAND with funding from the National Institute on Aging and the Social Security Administration.

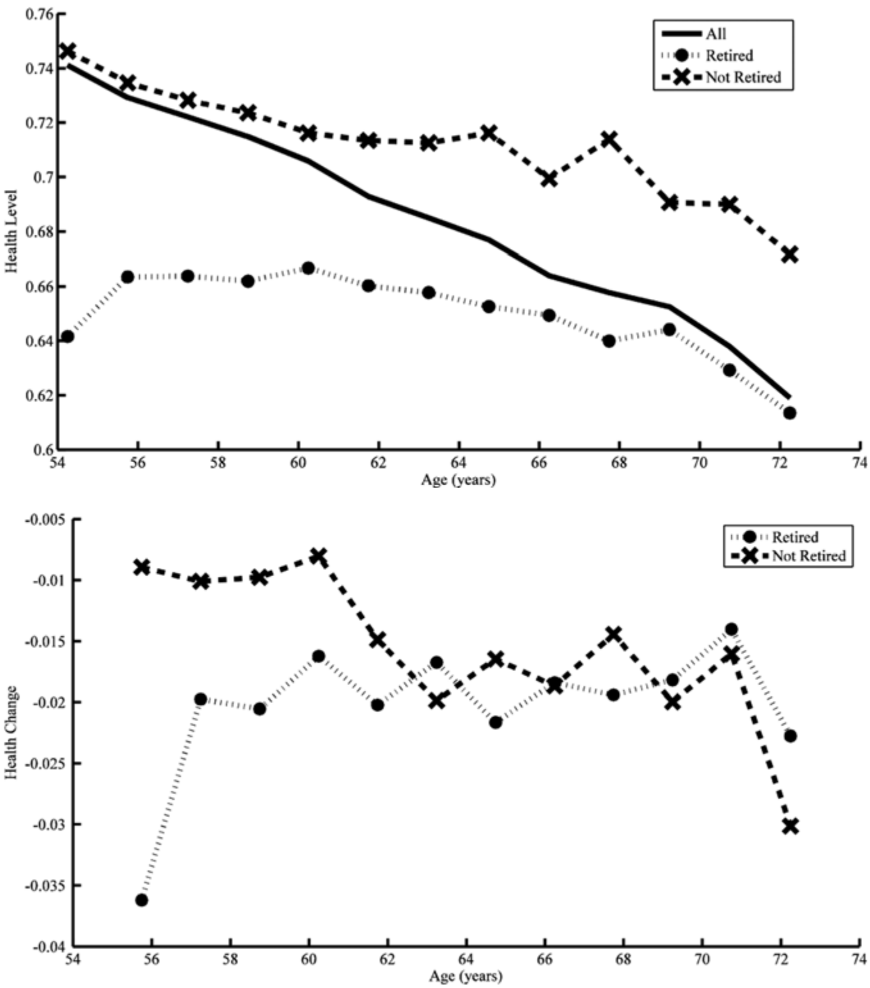

Figure 2

Average Health Movements

Notes: The first figure plots sample averages of health levels, grouped by age bins, conditional on retirement status. The second figure is a similar plot, but for health changes. Data are from the Health and Retirement Study (HRS). The reader may interpret general health as a continuous variable between zero and one, where zero denotes exceptionally poor health and one represents excellent health. Section III provides an extensive description of the health index.

comprehensive: In addition to the prior example of a spurious link between retirement and a concurrent large health decline, there remain many unobserved factors driving changes in health that may influence retirement decisions. Individuals might have some private beliefs about their health that impact their labor supply choices. A worker may forecast the onset of arthritis within five years and alter his or her retirement plans accordingly. Alternatively, an individual's planned retirement might coincide with a cancer diagnosis purely by chance. Whether expected or unexpected by individuals, these types of unobserved events would bias estimates of the retirement effect downward.

To correct such problems, this paper employs an instrumental variables strategy that demonstrates the effectiveness of subjective expectations variables, which have gained prominence in large survey questionnaires and analyses. The key instruments are individuals' predicted probability of working past ages 62 and 65, reported in the period they entered the sample.[7] In discussing the instrument's validity, it is helpful to conceptually divide the unobservable health factors that are correlated with retirement into two groups: factors that are anticipated by the individual and factors that cannot be anticipated. The instrument is orthogonal to unanticipated retirement-causing health changes by construction, because it was reported long before retirement occurred, but it may be correlated with individuals' private beliefs (or anticipations) regarding their health evolution. The strategy is to orthogonalize the instrument with respect to this anticipated component by proxying for it with a set of covariates: the respondents' parents' age (or age at death) as well as respondents' health level and health behavior characteristics observed at their period of entry into the panel. The intuition behind these proxies is that they combine respondents' historical belief-formation (original health information) with their future expectations (genetic factors that are captured by parents' age).

The primary conclusion is that retirement exerts a beneficial and statistically significant impact on individuals' future health prospects. The main estimate is interpreted as the local average treatment effect (LATE) of retirement on health change. Section VI clarifies the meaning of the LATE. Over an average length retirement spell (7.4 years in the sample), the effect is approximately equivalent to the prevention of "one-quarter" of one ailment condition, such as arthritis, or smaller fractions of more severe ailments. Additionally, the estimates are robust to three alternate specifications: a reestimation using a different definition of retirement, an examination of the model's predictions on various subsamples, and a comparison of the main estimates to those using alternate health indices which incorporate different weighting schemes or health information.

As a natural corollary to this question, it is helpful to explore possible channels through which retirement could influence health. Perhaps retirees alter their health-related behaviors, such as exercise, (less) smoking, or preventative care measures, following retirement. Section VI analyzes some stylized facts on smoking and exercise levels to explore possible direct effects of retirement on these health behaviors. These investigations yield evidence for an intuitive explanation of the main result: Retire-

7. The empirical work is restricted to the sample of individuals that enter the survey unretired. This is a relatively minor constraint, given that most individuals between the ages of 50–60 are still working.

ment may benefit health through behavioral channels, so that with additional leisure time, many retirees invest in their health via healthy habits.

The paper is organized as follows: Section II discusses related research; Section III describes the data, the specific sample used in the estimation, and the main health index; Section IV presents the econometric model; Section V discusses the estimation results from the main model and robustness checks; and Section VI proposes some interpretation of the results, analysis of health-related behaviors with respect to retirement, and an example of a simple policy analysis related to Social Security.

II. Related Literature

A few studies have sought to measure the impact of retirement on health. Dave, Rashad, and Spasojevic (2008) employed a fixed effects estimation strategy to control for time- invariant unobserved characteristics of individuals that are correlated with both retirement and health (these may include unobserved health issues, retirement preferences, or risk-taking behaviors). Their fixed effects estimates decreased relative to OLS estimates but did not switch sign, suggesting that retirement is harmful to health. In order to address the possibility of between-period retirement-causing health shocks, they performed the fixed effects estimation on a variety of sample stratifications. For example, they postulated that continuously insured individuals who do not report major health declines in the two previous periods are those least likely to experience a subsequent between-period decline. Their estimates on the subsamples decreased slightly but still did not change sign. The authors noted that the various subsamples of individuals are selected ones. Thus, it is unclear that such constrained retirement effect estimates are representative of the unrestricted sample. Additionally, even if individuals with better health histories are less likely to experience dramatic declines, some of them still do, so if estimates are biased, then they are certainly biased downward. While such specifications may be incomplete, the current study follows the insights of Dave, Rashad, and Spasojevic (2008) in estimating the instrumented model with individual fixed effects as well as examining similar sample stratifications as robustness checks.

Other papers have utilized instrumental variables strategies. Neuman (2008) devised an IV approach to tackle the issue of time-varying sources of endogeneity. His instrument set included spousal work-history and age dummies, variables regarding individuals' eligibility for private pensions, and binary indicators of age thresholds — 62 and 65 — the entitlement ages for Social Security and Medicare. Neuman found that retirement decreased the likelihood of a health decline, but his study faced a few limitations. His health change variables were loosely grouped binary encodings of whether individuals experienced a health decline since the previous period.[8] This technique did not incorporate the severities of ailments, differences in grouped ailments, how various ailments might respond differently to retirement, nor co-movements between the various health indicators. Additionally, the validity of spousal information and private

8. For instance, an occurrence of high blood pressure, stroke, or diabetes was coded as a general "health decline" for the "chronic conditions" health change indicator.

pension instruments is questionable if individuals can effectively predict their health evolution. For example, workers may have chosen a particular pension plan depending on how they expected their health to change. Their choices may have been correlated with spousal Social Security eligibility, compounding the issue. Overall, these instrumental variables are not as strong predictors of retirement behavior as directly reported retirement expectations. The instrument set in the current study demonstrates the power of subjective expectations variables to yield a strong first-stage regression and to permit a straightforward argument for validity.

Several others have studied the health consequences of retirement in various contexts. Bound and Waidmann (2007) estimated a beneficial retirement effect within the United Kingdom's ELSA data set, using features of the United Kingdom's public pension system as instrumental variables. Coe and Zamarro (2011) implemented an IV approach applied to Europe's SHARE data set using country-specific differences in retirement ages as instruments. They estimated a small positive retirement effect on self-reported health. Charles (2004) and Zhan et al. (2009) focused on psychological outcomes, also using discontinuous retirement incentive variables to uncover a positive influence. Rohwedder and Willis (2010) combined cross-country data from the HRS, ELSA, and SHARE to estimate a negative effect of early retirement on cognition. While it is not clear that findings based on European data should extend to the United States, literature suggests that, in general, retirement appears beneficial to future health outcomes, a result bolstered by the current study.

III. Data Description

A. Rand HRS 2010

The Health and Retirement Study (HRS) is a comprehensive biennial survey taken from 1992 through 2010. The survey's design and data collection have been organized by the University of Michigan and the National Institute on Aging. The Rand corporation has publicized a clean and user-friendly version of the data set. The original 1992 HRS cohort is a nationally representative sample of individuals born between 1931 and 1941 who reside in households. Survey respondents are re-interviewed every other year. The panel has added four more cohorts since its inception: AHEAD[9] (individuals born before 1924), Children of the Depression (CODA, born between 1924 and 1930), War Babies (WB, born between 1942 and 1947), and Early Baby Boomers (EBB, born 1948 to 1953).[10] Interviews have also been given to spouses of married or partnered respondents. The HRS utilizes a complex survey design that oversamples African-Americans, Hispanics, and Floridians (sampling weights, clustering, and strata variables are provided).

The questionnaire delves into an extensive set of topics: demographics, self-reported and doctor-diagnosed health characteristics, health insurance, finances, Social Security history, pension plans, retirement plans, and employment history. As a result, the HRS contains the proper ingredients to study the connection between retirement and

9. Stands for "The Study of Assets and Health Dynamics Among the Oldest Old."
10. The new cohorts were added in 1993, 1998, 1998, and 2004, respectively.

health. With ten waves of collected data, the survey has enough depth to effectively track the parameters of interest through time.

B. Construction of the Sample

The sample taken from the HRS in the current study contains the following restrictions: respondents must have worked for at least ten years, and they must have been employed during the period in which they entered the survey (this is due to the construction of the instrument set, to be described in Section IV). After eliminating respondents who do not meet these criteria, respondents in the AHEAD cohort (too old—the youngest are 68 in 1993), respondents who enter the sample under 50 years old, and respondents with missing values for race (eight individuals) and education (69 individuals), 10,632 individuals remain in the sample. The HRS includes several different questions about retirement status, most notably an hours-worked variable and a "completely" versus "partially" versus "not" retired indicator. Following the work of Gustman and Steinmeier (2000), individuals are considered retired if they report complete retirement or if they report partial retirement and work less than 20 hours per week on average. Individuals are listed as not retired if they report that they are not retired or if they report partial retirement along with at least 20 hours of work per week.[11] The sample excludes the 2,801 individuals who returned to the labor force from retirement. After omitting observations with missing values for key variables, the final sample size is 31,545 pooled observations for 6,276 distinct individuals. This provides an average longitudinal depth of five time periods per individual, or about ten years. Table 1 contains summary statistics of key variables in the pooled sample, conditional on labor force status. Table 2 contains summary statistics for key initial period variables including those that form the instrument set.

C. Health Measurement

Various indexing techniques are common in health literature, including weighting, factor analysis, and item response theory methods.[12] This paper adopts the first option; the primary health measure is derived from a straightforward weighting scheme. Robustness checks in Section V compare various alternate weighting schemes. Although omitted from the text, item response health indices did not produce substantially different results, either in their qualitative characteristics or their statistical significance. In general, the main health index is a weighted sum of "objective" doctor-diagnosed health variables and "subjective" self-reported health status.

A health index reduces relevant health-related information to a scalar value that represents general health. In order to use regression analysis to explain variation in health changes, it is necessary to assume that health is a unidimensional trait (because it is the dependent variable). The index is built from ten categorical variables. One is an ordinal response "subjective" self-reported health variable,[13] and the rest are

11. Section V includes a robustness check that alters the definition of retirement.
12. Jurges (2007); Lange and McKee (2011); McDowell (2006); Madrian, Mitchell, and Soldo (2007).
13. The corresponding question from the HRS questionnaire is: "Would you say your health is excellent, very good, good, fair, or poor?"

Table 1
Summary Statistics: Workers and Retirees

	Workers		Retirees	
	Mean	Standard Deviation	Mean	Standard Deviation
Self-reported health[a]	2.308	0.942	2.742	1.076
High blood pressure*	0.389	0.488	0.571	0.495
Diabetes*	0.109	0.312	0.201	0.401
Cancer*	0.065	0.247	0.155	0.362
Lung problems*	0.044	0.206	0.106	0.308
Heart problems*	0.111	0.314	0.244	0.429
Stroke*	0.014	0.117	0.055	0.227
Psychological problems*	0.098	0.298	0.165	0.371
Arthritis*	0.376	0.484	0.616	0.486
Obese*	0.283	0.450	0.298	0.458
Main health index[b]	0.803	0.151	0.692	0.190
Female*	0.432	0.495	0.466	0.499
Age (years)	59.4	4.26	66.4	5.47
Black*	0.075	0.264	0.085	0.279
Hispanic*	0.060	0.238	0.046	0.209
Married*	0.696	0.460	0.674	0.469
Assets (if > 0, in $)	489,759	1,712,087	532,497	1,276,347
Debt (if > 0, in $)	2,010	48,646	1,308	20,896
Less than high school*	0.033	0.178	0.058	0.233
Some high school*	0.062	0.241	0.105	0.307
High school diploma (or GED)*	0.296	0.456	0.352	0.478
Some college (or AA)*	0.257	0.437	0.228	0.419
Bachelor's degree*	0.158	0.365	0.113	0.316
Graduate degree*	0.194	0.395	0.145	0.352
Number of observations (pooled)	15,786		15,759	

Notes: *Binary indicator; estimates refer to sample proportions rather than sample averages.
a. 1 ~ Excellent health, 2 ~ very good, 3 ~ good, 4 ~ fair, 5 ~ poor.
b. See Section III.

"objective" binary response doctor-diagnosed health conditions. They include heart problems (heart attack, coronary heart disease, angina, or congestive heart failure), high blood pressure or hypertension, stroke or transient ischemic attack, diabetes or high blood sugar, chronic lung disease (aside from asthma), arthritis or rheumatism, cancer (aside from benign skin cancer), psychological problems (emotional,

Table 2
Summary Statistics: Initial Period Characteristics

	Mean	Standard Deviation
High blood pressure*	0.322	0.467
Diabetes*	0.078	0.268
Cancer*	0.041	0.198
Lung problems*	0.043	0.204
Heart problems*	0.090	0.286
Stroke*	0.007	0.085
Psychological problems*	0.068	0.251
Arthritis*	0.295	0.456
Obese*	0.252	0.434
Age (years)	55.1	3.04
Mother's age (if deceased)	69.6	14.07
Father's age (if deceased)	69.3	13.42
Mother's age (if living)	76.8	4.83
Father's age (if living)	79.4	4.98
Vigorous activity?[a]*	0.337	0.473
Smokes currently?[b]*	0.235	0.424
Self-reported probability of working past 62 (%)	51.0	38.8
Self-reported probability of working past 65 (%)	28.2	33.7
Number of observations	6,276	

Notes: *Binary indicator; estimates refer to sample proportions rather than averages.
a. Whether respondent engages in "vigorous physical activity" three or more times a week.
b. Whether respondent is "a current smoker."

nervous, or psychiatric problems), and obesity (indicated by a body-mass index[14] greater than 30).[15]

Bound et al. (1999) provide the framework for the construction of the main health index. They performed an ordered probit regression of self-reported health on a set of objective health characteristics. They then formed predictions of self-reported health for each individual based on the probit results, and those predictions became the final index. As all covariates were objective health characteristics, variation in the index was produced solely by individual differences in those objective health categories.

14. Calculated by dividing an individual's self-reported weight in kilograms by his or her self-reported height in meters squared.
15. The HRS questionnaire queries new interviewees: "Has a doctor ever told you that you have . . . " Repeated respondents are asked: "Since we last talked to you, that is since [last interview date], has a doctor told you that you have . . . "

Subjective self-reported health contributed only in the calculation of the probit coefficient estimates, thereby determining "weights" for each doctor-diagnosed condition. In the current study, the main model uses a modified technique to produce variation from both objective health conditions and self-reported health. The index is constructed as follows:

1. Estimate ten separate probit models, each one with a different health condition on the lefthand-side and the remaining nine health conditions on the righthand-side. (Note that this set of 10 includes all objective conditions as well as self-reported health.)
2. For each probit model and for each observation, generate a prediction of the dependent variable.
3. For each probit model, normalize the predictions to lie between zero and one, where outcomes closer to one indicate better health.
4. For each observation, average across all ten predictions to calculate the observation's final health index.

Such a procedure allows for an unprejudiced weighting scheme, as it is unclear how to otherwise integrate variation from both objective and subjective sources into the index. An alternate option would be to use a dependent variable that is not indicative of a distinct health characteristic in the "weighting-choice" probit model. Section V contains a robustness check using a variable regarding health-related work limitations on the lefthand side, as well as a check using a pure form of the Bound et al. (1999) index.

In developing a health index, it is important to consider the usefulness of including both "subjective" and "objective" survey information. On the one hand, there is evidence that "objective" variables are not, in fact, free of biases. Baker, Stabile, and Deri (2004) matched health-related Canadian survey data with respondents' official health records, finding strong evidence of both false negatives and false positives in the self-reported, "objective" ailment reports. This casts doubt on the idea that there is a clear distinction between "objective" and "subjective" indicators and implies that measurement error may enter the discussion (Section IVD further develops this issue). On the other hand, there is a clear distinction between the two types of health information based on survey-question wording: "Subjective" self-reported health (on a discrete scale from one to five) is very broad, while "objective" ailment queries are quite specific. Thus a natural concern is the extent to which including subjective indicators in the health index enhances its explanatory power: What, if anything, does subjective health add, after controlling for objective measures?[16]

Because a stated goal of this paper is to compute a health index that best represents "general health," such an index must incorporate as much relevant health-related information as possible. Many researchers in the health-labor field have estimated their results using objective and subjective health indicators separately, often finding differences that stem from the objective-subjective choice.[17] Indeed, the discrepancies themselves provide evidence that "subjective" measures contain *something* that

16. For instance, in implementing a version of the Bound et al. (1999) index, Coe and Zamarro (2011) refer to "multicollinearity problems that arise when including both objective and subjective measures of health as controls."

17. Anderson and Burkhauser (1985); Bound and Waidmann (2007); Neuman (2008); Dave, Rashad, and Spasojevic (2008).

objective measures do not. (The question of whether that *something* includes "good" health-related information or "bad" noise is left for Section IVD.) Others have employed health indices that simultaneously incorporate subjective and objective information.[18] Since such an index is utilized extensively below, it is also valuable to formulate a simple test to ensure that subjective indicators do indeed supplement the objective health information: Estimate a Bound et al. (1999)-style ordered probit with self-reported health regressed on the nine doctor-diagnosed ailment conditions, and then check various goodness-of-fit criteria to ensure that the self-health grades are not "too strongly" predicted. For the main HRS sample, some corresponding fit-metrics are McFadden's pseudo R^2 of 0.103, Cox & Snell's pseudo R^2 of 0.255, and an adjusted count R^2 of 0.102, which captures the predictive power on top of the baseline count R^2 of 0.418 from the intercept-only model. Thus there is strong evidence that doctor-diagnosed ailments provide some predictive power for self-reported health, but there is still a large component of those health scores that remains unexplained.

IV. Econometric Model

This section presents the theoretical foundation of the empirical work. The first step is to investigate the baseline ordinary least squares (OLS) model and its limitations. The derivation of the corrected form of the model follows.

A. Baseline Model (OLS)

In the following model, t is the survey period, i is the individual, and X_{it} is a set of exogenous controls that include age, years of education, gender, race, marital status, log of value of assets, and log of value of debt.[19] RS_{it} ("short-term" retirement) is a dummy variable indicating that individual i retired in period t (but not before). RL_{it} ("long-term" retirement) is equal to one only when i retired in period $t - 1$ or before. Thus the retirement effect has two components: RS_{it} measures the short-term effect of recent retirement that occurred since the previous survey, and RL_{it} gauges the cumulative effect of retirement spells that are at least one time period (or two years) long. Specifications that further discretize the cumulative effect require additional lags of retirement, constraining the sample significantly.[20] ΔH_{it} is the change in the health index. A health change value less than zero corresponds to a decline in health. μ_{it} includes all unobserved factors that drive changes in health. The baseline specification is:[21]

18. Lange and McKee (2011); McHorney and Cohen (2000); Madrian, Mitchell, and Soldo (2007); Ware et al. (1995).

19. Log of debt and log of assets are conditional on debt and assets being greater than zero, respectively. In other words, log of debt (or assets) is set equal to zero if the observation's debt (or assets) level is less than one dollar.

20. In the case of a three-period discretization, a minimum of four observations would be required: four time periods permit three observed health changes, one for each of the necessary three observations of retirement status. Such a restriction would eliminate approximately one quarter of the sample.

21. Versions of the OLS and IV (developed in Section IVB below) models including period $t - 2$ health level indicators (self-reported health and each of the nine objective ailment conditions) as additional controls yielded only small changes in the results. Some coefficient estimates lost precision because the constrained sample was about 25 percent smaller. These results are available upon request.

(1) $\Delta H_{it} = \beta X_{it} + \theta_1 RS_{it} + \theta_2 RL_{it} + \mu_{it}$

The dependent variable is health change instead of health level because postretirement health changes are not susceptible to the simultaneity problem (retirement-causing health declines are no longer an issue because retirement has already occurred). However, retirement is still endogenous. RS_{it} and RL_{it} are correlated with μ_{it} because an individual's retirement decision may depend on health-related events that are unobserved (to the econometrician), biasing coefficient estimates. Estimates of θ_1 will be biased downward because retirement-causing health shocks between periods t and $t-1$ will be picked up by the RS_{it} indicator. OLS estimates of θ_2 may also suffer downward bias due to anticipated health declines that may be correlated with retirement.[22]

B. Corrected Model (IV)

An instrumental variables strategy aims to account for endogeneity in retirement. It is helpful to split the error term μ_{it} into three pieces. Let f_i represent fixed (by individual) unobserved heterogeneity correlated with health change, let a_{it} include time-variant unobserved health effects that are anticipated by the individual, and let u_{it} include all other (unanticipated) time-variant unobserved effects:

(2) $\mu_{it} = f_i + a_{it} + \mu_{it}$

a_{it} may be correlated with retirement because it encompasses hidden health characteristics (in particular, health expectations) that are naturally tied to the retirement decision. For instance, a worker might hold private knowledge about a hereditary heart condition that compels him or her to retire before its actual onset. u_{it} may be correlated with retirement due to unanticipated between-period health shocks influencing an individual's retirement decision. Lastly, f_i contains any possible time-invariant unobservables; retirement-related hidden anticipations could also lie in f_i.

The instrument set is based on two key questions from the HRS questionnaire: "What do you think the chances are that you will be working full-time after you reach age 62?" (There is an analogous question for age 65.) Because the sample includes only individuals who entered the survey while employed, these inquiries provide valuable information about their retirement preferences and expectations. Their responses are (by construction) uncorrelated with unanticipated retirement-causing health shocks u_{it}, because they are taken from individuals' initial observations, which preceded their retirement.

The retirement expectations instruments are likely correlated with a_{it}, so the next step is to orthogonalize them to a set of proxies for a_{it}. As with the instruments, the proxy variables are taken from each individual's initial period of observation. They include parents' age P_{i0}, the respondent's initial age AGE_{i0}, original-period health level indicators HI_{i0}, and original-period health behaviors HB_{i0}.[23] The IV strategy relies on

22. For instance, bias could be caused by diagnoses of ailments that are degenerative. Or, an individual may be aware of a predisposition for cancer and plan his or her retirement accordingly. Bias could also come through health events that occur very close to the survey date that impact labor supply choices.

23. In the regressions, P_{i0} is split into four variables: mother's age at death (or equal to zero if still living), mother's current age if still living (or equal to zero if deceased), and likewise for the father's status. HI_{i0} includes dummies for self-reported health and for the nine objective ailment conditions, and HB_{i0} contains the smoking and exercise dummies shown in Table 2.

the assumption that the residuals from the following two regressions are orthogonal to a_{it}:[24]

(3) $\Pr(\textit{Working past } 62)_{i0} = \psi_1 P_{i0} + \gamma_1 AGE_{i0} + \delta_1 HI_{i0} + \lambda_1 HB_{i0} + \varepsilon_{1,i0}$

(4) $\Pr(\textit{Working past } 65)_{i0} = \psi_2 P_{i0} + \gamma_2 AGE_{i0} + \delta_2 HI_{i0} + \lambda_2 HB_{i0} + \varepsilon_{2,i0}$

For this assumption to hold, the set of proxies must be correlated with the component of retirement expectations that is linked to future health change. In other words, the following must hold (where $\hat{\varepsilon}_{1,i0}$ and $\hat{\varepsilon}_{2,i0}$ are the residuals):

(5) $corr(\{\hat{\varepsilon}_{1,i0}, \hat{\varepsilon}_{2,i0}\}, a_{it}) = 0$

The next subsection discusses this assumption in more detail.

The last step in constructing the instrument set is to generate interactions of the residuals with binary indicators of whether the individual is under age 62 or is age 62–65 (to reflect discontinuous retirement incentives at these ages due to Social Security and Medicare eligibility). Thus the instrument set is six-dimensional:

1. Under-62 dummy
2. Age 62–65 dummy
3. $\hat{\varepsilon}_{1,i0}$
4. $\hat{\varepsilon}_{2,i0}$
5. $\hat{\varepsilon}_{1,i0} \times$ Under-62 dummy
6. $\hat{\varepsilon}_{2,i0} \times$ Age 62–65 dummy

The interaction terms function as slope-differentials for the dummies, and the instrument set as a whole yields a strong first-stage regression. It is overidentified, which permits validity tests of the corrected model (results in Section V). The dummy variable interactions also ensure that the residual component is time variant, allowing fixed effects specifications to be used to address potential endogeneity stemming from f_i.

The final model utilizes two stage least squares (2SLS) to estimate Equation 1, instrumenting endogenous variables RS_{it} and RL_{it} with the variables described above. 2SLS yields consistent estimates of the regression coefficients under the assumptions previously discussed and in the next subsection.

C. Instrument Validity

In order for estimates of θ_1 and θ_2 to be consistent, the standard IV assumptions must be satisfied. The first-stage regressions imply that the instrument set strongly predicts retirement in each period.[25] The crucial assumption is that the instruments must be uncorrelated with μ_{it}, which can be decomposed into the three components seen in Equation 2. The instruments are orthogonal to u_{it} by construction, and fixed-effects can circumvent possible correlation with f_i. However, the instruments must also be uncorrelated with a_{it}. Recall that the instrument set has two components: residuals calculated from orthogonalization Equations 3 and 4 as well as age dummies. Exogeneity of the age dummies follows because individuals are not predisposed to have certain

24. Results from these regressions are available upon request.
25. See Section V for details.

health shocks at those ages versus any other similar age. They are linked to retirement because 62 and 65 are the entitlement ages for Social Security and Medicare benefits. Thus it remains only to argue that Equation 5 holds. The retirement expectations instruments—Pr(*Working past* 62)$_{i0}$ and Pr(*Working past* 65)$_{i0}$—contain two main pieces of information about an individual:

1. Information on retirement expectations
2. Information on retirement preferences

Individuals' retirement preferences contain only exogenous variation that is correlated with their actual retirement. Individuals' retirement expectations may be endogenous because they are tied to their health expectations (namely, both the instruments and a_{it} contain this information). The hypothesis is that individuals form their hidden (to the econometrician) retirement-related health expectations based on hereditary health trends (proxied by parents' age) and past health history (proxied by initial-period health levels, behaviors, and age). Thus the "expectations component" is removed from the orthogonalized instruments $\hat{\varepsilon}_{1,i0}$ and $\hat{\varepsilon}_{2,i0}$, leaving only exogenous variation in retirement preferences. If the set of proxies omits crucial information that is correlated with a_{it}, then estimates may be biased downward due to anticipated retirement-causing health shocks.[26] In summary, instrument validity relies on the assumption that the proxies are robust enough to predict anticipated effects. If this holds, 2SLS estimates of the regression parameters are consistent.

It is reasonable to suspect that parents' age and initial health are inadequate proxies for the component of expected health that is contained in the retirement expectations variables. In this regard, it is helpful to estimate IV models using the "uncleaned" instruments, Pr(*Working past* 62)$_{i0}$ and Pr(*Working past* 65)$_{i0}$, in place of the "cleaned" residuals, $\hat{\varepsilon}_{1,i0}$ and $\hat{\varepsilon}_{2,i0}$, in the instrument set. The discussion below omits these estimation results for brevity, but it is important to note that they do not differ significantly from the main results.[27] "Uncleaned" instruments yield IV estimates of θ that are slightly less statistically significant, but corresponding Sargan-Hansen *J*-statistics are indistinguishable from those of the main model. This suggests that the proxies are not strong predictors of expected health, but at the same time, "improperly cleaned" expected health information does not seem to adversely affect instrument validity. Thus it may simply be that individuals are not able to effectively forecast their future retirement based on their expected health evolution.

D. Measurement Error

Measurement error can be an issue when working with health indices. If it is present, then true health, h_{it}, is unobserved. In this case, the econometrician observes:

$$(6) \quad H_{it} = h_{it} + \eta_{it}$$

where η_{it} is measurement error. There are two notable potential sources of measurement error. First, the health index may not capture all information that describes one's

26. One can imagine less plausible stories with opposite bias, such as an individual who anticipates a health increase and subsequently retires in order take advantage of his newfound health.

27. Results are available upon request.

health. In an ideal setting, the data would contain a complete set of health-related information such as cholesterol levels, blood and liver tests, nutrition, cardiovascular status, and an extensive disease history. All missing information is contained in η_{it}. However, the information included in the health index should be more descriptive of general health than such missing characteristics. For example, an individual who has a resting heart rate of 65 beats per minute may be healthier than one with a heart rate of 85 (all else equal), but pulse rate is not as consequential as doctor-diagnosed hypertension. These types of omissions may yield only classical measurement error in the dependent variable, thus inflating standard errors. Even if this source of measurement error is not purely classical, any potential correlations between the error term and explanatory variables should be negligible because the health index contains enough crucial health-related information. For instance, a worker is less likely to retire on account of minor dental problems than he or she is after experiencing a heart attack. In any case, these types of issues would be corrected via the IV specification.

A second source of measurement error is known as "justification bias," which refers to retirees' tendencies to exaggerate their poor health in order to provide socially acceptable justification for their retirement. This phenomenon has been studied extensively by Bazzoli (1985), McGarry (2004), and others. Under such misreports, observed health would be understated for retirees, meaning that η_{it} may be correlated with RS_{it}, RL_{it}, or both. It is possible to show that the bias would work against the conclusion that retirement preserves health (thus making it harder to obtain significant positive estimates of θ_1 and θ_2).[28] If η_{it} represents justification bias, it has the following form:

(7) $$\eta_{it} \begin{cases} = 0 & \text{if } i \text{ is not retired in period } t \\ \leq 0 & \text{if } i \text{ is retired in period } t \end{cases}$$

Consider the following simplified version of the structural Equation 1:

(8) $$\Delta H_{it} = \theta_1 RS_{it} + \theta_2 RL_{it} + \mu_{it} + \Delta\eta_{it}$$

The size of the bias may be constant once an individual retires. In this case, $\Delta\eta_{it}$ is correlated with RS_{it} but not with RL_{it}. In either case, since the retirement expectations instrument is correlated with retirement, it is also correlated with $\Delta\eta_{it}$. The next section shows that the final estimate of RS_{it}'s coefficient is zero and the final estimate of RL_{it}'s coefficient switches signs (negative to positive - going from OLS to 2SLS). Therefore, even if justification bias does affect the corrected estimate of θ_2, the 2SLS estimate serves as a lower bound for θ_2 (and it is, at worst, masking a true positive value for θ_1). In other words, justification bias may act against the sign change, but the sign switches nevertheless. Note that the main source of justification bias in the health index should be self-reported health. Given questionnaire wording, respondents are less likely to exaggerate their responses to the objective doctor-diagnosed conditions (but they still may, as mentioned above in reference to the work of Baker, Stabile, and

28. There could also be an opposite bias: One can imagine a "role bias" in which once individuals enter into retirement they may feel healthier than they did while working because their role in retirement is less physically or mentally demanding. In this case, observed health would be biased upward following retirement, falsely inflating retirement's health-preserving effect. A robustness check in Section VB further develops this notion.

Deri 2004). As an additional test, the next section includes estimations of the model with a health index using only those objective conditions.

V. Main Results and Robustness Checks

This section reports the main empirical findings and describes three robustness checks: a reestimation using a different definition of retirement, an examination of the model's predictions on various subsamples, and a comparison of the main estimates to those using alternate health indices which incorporate different weighting schemes or heath information.

A. Main Results

Table 3 displays four estimations of Equation 1:

1. Baseline model estimated by pooled OLS
2. Baseline model estimated via fixed effects (FE) regression
3. Corrected model estimated via random effects (RE) regression
4. Corrected model estimated via fixed effects regression

In the baseline model (Model 1 in the table), the exogenous controls behave as expected. Women tend to experience less severe health changes relative to men. The health change index has a standard deviation of 0.143. Thus womens' health changes are on average less than 1 percent of a standard deviation better than mens' (across the two-year periods of observation). Individuals identified as black and Hispanic experience less severe health changes on average, but the effect is also very small (about 2 percent of a standard deviation). Wealth is correlated with health preservation, and education-level dummy estimates increase with more years of education. For instance, an individual with a bachelor's degree tends to experience a more favorable health change (by about 6 percent of a standard deviation in the health index), compared to an individual with zero years of high school. Table 3 presents baseline fixed effects estimates in the Model 2 column. Fixed effects only identifies coefficients for time-variant characteristics, whose estimates are very similar to those in Model 1. The notable changes are in the age polynomial estimates, which switch signs but still fail to gain statistical significance, and the "is married" indicator, which is now negatively associated with health changes by about the same magnitude as the education dummies.

In the baseline models, estimates of retirement effects are predominantly negative and statistically significant. In Model 1, long-term retirement is insignificant but associated with an average decline of 0.000148 in the health index, while contemporaneous (short-term) retirement is linked to a much stronger and significant decline. The mechanical interpretation of RS_{it}'s coefficient estimate is to say that a "new retirement" (occurring between the current and previous surveys) is associated with a decline of 0.0137 in the health index, holding all other observable characteristics fixed. The corresponding interpretation for RL_{it} is that an average-length retirement spell (7.4 years, conditional on retirement having occurred two or more years before the latest survey) is associated with an average decline of 0.000148 in the health index. Section VI presents some more practical interpretations of these results. Fixed effects

Table 3
Main Regression Results

Dependent Variable: ΔH_{it}	1	2	3	4
Female	0.00124*		0.00137	
	(0.000732)		(0.00108)	
Age	0.0000228	–0.000856	–0.000508	–0.00478*
	(0.00135)	(0.00168)	(0.00218)	(0.00281)
Age^2	–0.0000000	0.00000316	0.00000487	0.0000284
	(0.0000105)	(0.0000130)	(0.0000168)	0.0000199)
Black	0.00345***		0.00345**	
	(0.00115)		(0.00158)	
Hispanic	0.00402**		0.00394*	
	(0.00159)		(0.00215)	
Married	–0.000152	–0.00712**	–0.0000438	–0.00743**
	(0.000908)	(0.00313)	(0.00122)	(0.00314)
log(assets)	0.00118***	0.000879	0.00116***	0.000794
	(0.000251)	(0.000584)	(0.000269)	(0.000595)
log(debt)	–0.000195	–0.000740	–0.000236	–0.000853
	(0.000473)	(0.000763)	(0.000447)	(0.000779)
Some high school	0.00565***		0.00569**	
	(0.00206)		(0.00260)	
High school diploma	0.00684***		0.00680***	
	(0.00184)		(0.00232)	

(continued)

Table 3 (*continued*)

Dependent Variable: ΔH_{it}	1	2	3	4
Some college	0.00584***		0.00575**	
	(0.00192)		(0.00243)	
Bachelor's degree	0.00893***		0.00886***	
	(0.00203)		(0.00268)	
Graduate degree	0.00895***		0.00893***	
	(0.00200)		(0.00261)	
Constant	–0.0455	0.0109	–0.0316	
	(0.0431)	(0.0537)	(0.0699)	
RL_{it}: Long-term retirement	–0.000148	0.00544***	–0.00223	0.0194*
	(0.00102)	(0.00207)	(0.00446)	(0.0101)
RS_{it}: Short-term retirement	–0.0137***	–0.00857***	–0.00827	0.0177
	(0.00204)	(0.00233)	(0.0127)	(0.0161)
Number of observations (pooled):	31,545	31,545	31,545	31,545

Notes: Statistical significance is indicated by p-values: *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered by individual. The reference category for education dummy variables is "zero years of high school." Model 1 refers to the baseline model estimated via pooled OLS. Model 2 is the baseline via fixed effects. Model 3 is the corrected IV model via random effects, and Model 4 is the corrected IV model via fixed effects.

estimates change notably to 0.00544 and –0.00857, respectively. Following the work of Dave, Rashad, and Spasojevic (2008), short-term retirement is still negative since fixed effects do not cleanse estimates of between-period retirement-causing health shocks. The long-term retirement variable has switched sign but may still suffer from endogeneity bias, so the next step is to consider the IV specifications.

Table 4 contains the first stage regression results for the corrected models (Models 3 and 4 in Table 3). The dependent variables are RS_{it} and RL_{it}, and the covariates include all exogenous regressors as well as the instruments excluded from the structural equation. The instrumental variables are strong predictors of retirement. F-statistics from joint significance tests of the instruments show that first-stage regressions for long-term retirement are much stronger, although still at acceptable levels for short-term retirement in both RE and FE specifications. To interpret the instruments' coefficient estimates, the residuals ($\hat{\varepsilon}_{1,i0}$ and $\hat{\varepsilon}_{2,i0}$) may be viewed as variables that are strongly and positively correlated with the expected retirement indicators, Pr(*Working past* 62)$_{i0}$ and Pr(*Working past* 65)$_{i0}$. For instance in Column 1, for an individual under 62, an additional percentage point in predicted probability of working past 62 corresponds to a –0.00229 + 0.00215 = –0.00014 (smaller) probability of being currently "long-term retired." In Column 2, the same exercise yields a –0.000423 – 0.00144 ≈ –0.001 change in probability of being "short-term retired."[29] The coefficient interpretations are qualitatively similar for the random effects model's "working past 65" instruments, as well. In Columns 3 and 4, estimates for $\hat{\varepsilon}_{1,i0}$ and $\hat{\varepsilon}_{2,i0}$ are excluded due to the fixed effects specification, so it becomes more difficult to naturally interpret their interactions' coefficients since the net effect depends on the unidentified time-invariant fixed effect. Thus the random effects first stage regression is more easily interpretable and confirms that the instruments are correlated to retirement in a logical manner.

Models 3 and 4 in Table 3 display the results from the IV specifications. Using random effects, the estimates of current period retirement and long-term retirement go to zero. A Sargan-Hansen test for Model 3 yields a J-statistic of 9.988 (p-value of 0.0406), rejecting that the instruments are exogenous. The IV fixed effects specification, however, has a J-statistic of 0.523 (p-value of 0.7698), suggesting that one must account for time-invariant effects to attain instrument validity. Model 4 thus is the "fully corrected" model. The long-term retirement coefficient ($\hat{\theta}_2 = 0.0194$) switches sign and can be interpreted as the local average treatment effect (LATE) of retirement on health change. Section VI discusses this in more detail. As one might expect, the "cumulative retirement" effect is much stronger than short-term retirement, whose coefficient estimate goes to zero in the final model. Identifiable exogenous variables (age, assets, and marital status) do not substantially change across the four specifications. The next three subsections test the robustness of these results.

B. Robustness Check: Alternate Definition of Retirement

One notable point of flexibility in the econometric specification is how to define retirement. This is a central characteristic that is also related to the channels through which retirement may act upon health. For some individuals, retirement may simply refer to

29. Note that these figures are approximate because the residuals are not perfectly correlated with the predicted probabilities.

Table 4
First Stage IV Results

Dependent Variable	RL_{it} (RE)	RS_{it} (RE)	RL_{it} (FE)	RS_{it} (FE)
Female	0.0222***	–0.00782**		
	(0.00704)	(0.00347)		
Age	0.0119*	0.0817***	0.0176**	0.0847***
	(0.00641)	(0.00598)	(0.00805)	(0.00621)
Age^2	0.000176***	–0.000643***	0.000138**	–0.000669***
	(0.0000481)	(0.0000448)	(0.0000613)	(0.0000458)
Black	–0.00515	0.00375		
	(0.0102)	(0.00511)		
Hispanic	–0.0307**	0.00138		
	(0.0139)	(0.00694)		
Married	0.0132**	–0.00891**	0.0208*	0.00227
	(0.00638)	(0.00393)	(0.0124)	(0.0107)
Log(assets)	–0.00790***	0.00157*	–0.00592***	0.00572***
	(0.00123)	(0.000868)	(0.00174)	(0.00169)
Log(debt)	–0.0104***	0.00368**	–0.00888***	0.00854***
	(0.00173)	(0.00144)	(0.00219)	(0.00226)
Some high school	–0.0123	–0.00747		
	(0.0174)	(0.00842)		
High school diploma	–0.0369**	–0.00295		
	(0.0154)	(0.00751)		
Some college	–0.0423***	–0.00255		
	(0.0159)	(0.00785)		

Bachelor's degree	–0.0488***	–0.00699		
	(0.0174)	(0.00864)		
Graduate degree	–0.0408**	–0.0137		
	(0.0170)	(0.00840)		
Under age 62	–0.231***	–0.0000872	–0.224***	–0.00387
	(0.00814)	(0.00802)	(0.0107)	(0.00952)
Age 62–65	–0.164***	0.0753***	–0.160***	0.0723***
	(0.00609)	(0.00608)	(0.00799)	(0.00796)
$\hat{\varepsilon}_{1,i0}$	–0.00229***	0.000423***		
	(0.000131)	(0.0000678)		
$\hat{\varepsilon}_{2,i0}$	–0.00127***	0.000165**		
	(0.000141)	(0.0000714)		
$\hat{\varepsilon}_{1,i0}$ × Under–62 dummy	0.00215***	–0.00144***	0.00227***	–0.00175***
	(0.000107)	(0.0000899)	(0.000181)	(0.000138)
$\hat{\varepsilon}_{2,i0}$ × Age 62–65 dummy	0.000122	–0.00164***	0.000196	–0.00168***
	(0.000140)	(0.000137)	(0.000176)	(0.000184)
Number of observations (pooled):	31,545	31,545	31,545	31,545
F-statistic:	2,016.92	596.34	173.49	85.47

Notes: Statistical significance is indicated by p-values: : $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered by individual. The "RE" columns refer to first stage results for Model 3 in Table 3, and the "FE" models refer to first stage results for Model 4 in Table 3. Coefficient estimates of the constant term have been omitted for brevity, but were present in the estimations of all four models. The reference category for education dummy variables is "zero years of high school."

their exit from the labor force. For others, it could refer to fewer hours on the same job or in the same career, or perhaps the opportunity to begin a new part-time career. The retirement indicator used in the main model was meant to accommodate this array of possibilities. However, if the fully corrected IV model truly captures a strong retirement effect on health, the effect should also be present under alternate definitions of retirement.

This test reestimates Models 1–4 using a new definition. In each wave, survey-takers respond to the question: “Are you currently working for pay?” This binary indicator forms a simple retirement dummy. Table 5 contains regression results. Estimates are qualitatively very similar those from the main model. Comparing baseline RE to corrected FE, the long-term retirement coefficient estimate goes from zero to a positive value, while the short-term retirement estimate goes from a negative value to zero. Using the alternate indicator, retirement estimates tend to be larger in absolute value. Various conjectures could explain the larger estimates: Complete nonwork may provide retirees with even more opportunity for healthy practices compared to partial retirees. Or, the set of “full retirees” may contain a larger proportion of individuals who were involuntarily forced out of the labor force due to injury or illness. In general, the main finding that retirement drives positive health changes appears intact.

C. Robustness Check: Estimation on Subsamples

The next robustness check reestimates the various models on two different subsamples. Dave, Rashad, and Spasojevic (2008) suggested that healthier respondents and younger respondents should be less likely to experience sudden and severe illnesses leading to involuntary retirement. Under their hypothesis, retirement estimates in the baseline models should not be “as endogenous” as those from the main sample. Thus the relative change of baseline estimates to corrected ones should not be as large as in the main sample.

Table 6 contains retirement coefficient estimates from estimations restricted to these subgroups (the top portion of the table reproduces the main sample results for easy comparison). The younger subsample consists of only those survey respondents who entered the panel under age 58. The healthier subsample consists of only the individuals who were in the top 75 percent of the initial-period health index distribution. Comparing fixed effects models (Model 2 to Model 4), the younger sample’s final estimates are insignificant, but they increase (going from 0.00597 to 0.0146) by less than the main sample’s estimates (0.00544 to 0.0194). The same is true of the healthier sample, but statistical significance is maintained in this case.[30] These tests imply that the corrective measures perform as intended on reasonable stratifications of the sample.

D. Robustness Check: Alternate Health Indices

The final test performs estimations using two alternate health indices. The new health indices stem from similar calculations as the main health index detailed in

30. Note that the differences between all three models are not perceptible at reasonable levels of statistical significance.

Table 5
Robustness Check: Alternate Retirement Definition

Dependent Variable: ΔH_{it}	1	2	3	4
Female	0.00122*		0.00140	
	(0.000734)		(0.00109)	
Age	–0.000718	–0.000939	–0.00101	–0.00514*
	(0.00131)	(0.00165)	(0.00218)	(0.00273)
Age^2	0.00000528	0.00000331	0.00000856	0.0000286
	(0.0000102)	(0.0000128)	0.0000169)	(0.0000192)
Black	0.00387***		0.00389**	
	(0.00116)		(0.00159)	
Hispanic	0.00329**		0.00320	
	(0.00159)		(0.00215)	
Married	–0.000578	–0.00654**	–0.000475	–0.00654**
	(0.000915)	(0.00312)	(0.00122)	(0.00315)
Log(assets)	0.00112***	0.000608	0.00109***	0.000544
	(0.000248)	(0.000585)	(0.000269)	(0.000602)
Log(debt)	–0.000380	–0.00102	–0.000430	–0.00116
	(0.000473)	(0.000763)	(0.000450)	(0.000786)
Some high school	0.00453**		0.00446*	
	(0.00207)		(0.00261)	
High school diploma	0.00583***		0.00562**	
	(0.00184)		(0.00234)	

(*continued*)

Table 5 (*continued*)

Dependent Variable: ΔH_{it}	1	2	3	4
Some college	0.00469**		0.00440*	
	(0.00191)		(0.00246)	
Bachelor's degree	0.00720***		0.00686**	
	(0.00202)		(0.00272)	
Graduate degree	0.00729***		0.00697***	
	(0.00199)		(0.00268)	
Constant	–0.0185	0.0180	–0.0123	
	(0.0420)	(0.0525)	(0.0693)	
RL_{it}: Long-term retirement	0.000917	0.00781***	–0.00243	0.0283**
	(0.000991)	(0.00211)	(0.00503)	(0.0133)
RS_{it}: Short-term retirement	–0.0164***	–0.0101***	–0.0132	0.0380
	(0.00218)	(0.00249)	(0.0175)	(0.0249)
Number of observations (pooled):	31,491	31,491	31,491	31,491

Notes: Statistical significance is indicated by p-values: : $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors are clustered by individual. The reference category for education dummy variables is "zero years of high school." Model 1 refers to the baseline model estimated via pooled OLS. Model 2 is the baseline via fixed effects. Model 3 is the corrected model via random effects, and Model 4 is the corrected model via fixed effects.

Table 6
Robustness Check: Comparison to Subsamples

Dependent variable: ΔH_{it}	1	2	3	4
Main results (reproduced from Table 3)				
RL_{it}: Long-term retirement	–0.000148	0.00544***	–0.00223	0.0194*
	(0.00102)	(0.00207)	(0.00446)	(0.0101)
RS_{it}: Short-term retirement	–0.0137***	–0.00857***	–0.00827	0.0177
	(0.00204)	(0.00233)	(0.0127)	(0.0161)
Number of observations (pooled):	31,545	31,545	31,545	31,545
Younger subsample (initially under age 58)				
RL_{it}: Long-term retirement	–0.000382	0.00597***	–0.00404	0.0146
	(0.00113)	(0.00229)	(0.00505)	(0.0119)
RS_{it}: Short-term retirement	–0.0152***	–0.00927***	–0.00312	0.0201
	(0.00224)	(0.00253)	(0.0150)	(0.0179)
Number of observations (pooled):	25,848	25,848	25,848	25,848
Healthier subsample (initially top 75 percent of health distribution)				
RL_{it}: Long-term retirement	–0.00148	0.00364*	–0.00209	0.0169*
	(0.00105)	(0.00210)	(0.00452)	(0.0100)
RS_{it}: Short-term retirement	–0.0140***	–0.00909***	–0.00255	0.0209
	(0.00212)	(0.00239)	(0.0130)	(0.0163)
Number of observations (pooled):	28,889	28,889	28,889	28,889

Notes: Statistical significance is indicated by p-values: : $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered by individual. Coefficient estimates of exogenous controls (X_{it} in Equation 1) have been omitted for brevity, but were present in the estimations of each model. Model 1 refers to the baseline model estimated via pooled OLS. Model 2 is the baseline via fixed effects. Model 3 is the corrected model via random effects, and Model 4 is the corrected model via fixed effects. The main results are reproduced from Table 3 for ease of comparison to the two subsamples. The younger subsample consists of only those survey respondents who entered the panel under age 58. The healthier subsample consists of only the individuals who were in the top 75 percent of the initial-period health index distribution.

Section III. The first index is similar to that of Bound et al. (1999) and is calculated as follows:

1. Estimate an ordered probit model with a self-reported health on the lefthand-side and the remaining nine "objective" doctor-diagnosed conditions on the righthand-side.
2. Generate predictions of the self-reported health observations from the probit estimation.
3. Normalize the predictions to lie between zero and one, where outcomes closer to one indicate better health.

The variation in the index comes only from individuals' responses to the objective health questions (although it is scaled by their relation to self-reported health). Table 7 contains retirement estimates of the four econometric specifications using this index (it reproduces the main results in the top panel). Baseline estimates (Models 1 and 2) using the Bound et al. (1999) index are qualitatively similar to the estimate of the main model, although FE causes the coefficient estimate for RL_{it} to go to zero. The fully corrected model yields statistically insignificant results that are negative.

Other studies have encountered this (lack of significance) issue when using purely objective measures, including Neuman (2008) and Bound and Waidmann (2007). A possible explanation is due to a "role bias": Retirees may feel healthier than they did while working because their role in retirement is less physically or mentally demanding. In this situation, subjective self-reported health may be inflated because retirees' "perceived health" improves, even without any changes in "real health," which would be reflected in objective health indicators. This would bias estimates of retirement's effect upwards in the main model, explaining why the Bound et al. (1999) (the ailment-condition-only) index does not capture an effect. An alternate explanation, as discussed in Section IIIC, is that objective measures simply do not possess enough information to capture retirement's effect under the IV estimation strategy. There is strong evidence, from both related literature and simple empirical exercises described in Section IIIC, that subjective health variables enhance health measurement. Such evidence refutes the "role bias hypothesis" by indicating that it is the loss of information that causes the loss of significance, not the loss of an upward bias.

Since it may be more appropriate to use health indices that employ both objective and subjective health characteristics, the main results should be robust to alternative indexing schemes using both types of information. The third panel of Table 7 presents reestimations of the four models using a jointly objective and subjective health index, again derived from the Bound et al. (1999) indexing technique, with two differences: The lefthand-side variable of the probit model is a binary indicator of whether "health limits [the respondent's] ability to work" and the righthand-side variables now include self-reported health dummies in addition to the nine doctor-diagnosed conditions. Results are qualitatively similar to the main results with the exception that the short-term retirement estimate is now significant at the 5 percent level. Estimates using this index tend to be larger in absolute value than the main estimates, but it is not safe to compare across the two models since they have different dependent variables.

Overall, the main results appear robust to the choice of health index, as long as it

Table 7
Robustness Check: Alternate Health Indices

Dependent variable: ΔH_{it}	1	2	3	4
Main results (reproduced from Table 3)				
RL_{it}: Long-term retirement	–0.000148	0.00544***	–0.00223	0.0194*
	(0.00102)	(0.00207)	(0.00446)	(0.0101)
RS_{it}: Short-term retirement	–0.0137***	–0.00857***	–0.00827	0.0177
	(0.00204)	(0.00233)	(0.0127)	(0.0161)
Number of observations (pooled):	31,545	31,545	31,545	31,545
Bound et. al (1999) (self-reported health weighted) index				
RL_{it}: Long-term retirement	–0.00395***	–0.00236	–0.00354	–0.00414
	(0.000868)	(0.00155)	(0.00293)	(0.00689)
RS_{it}: Short-term retirement	–0.0114***	–0.00863***	–0.00621	–0.00361
	(0.00138)	(0.00157)	(0.00834)	(0.0112)
Number of observations (pooled):	31,545	31,545	31,545	31,545
Health-limiting weighted index				
RL_{it}: Long-term retirement	–0.000782	0.00905***	0.00138	0.0399**
	(0.00160)	(0.00320)	(0.00713)	(0.0159)
RS_{it}: Short-term retirement	–0.0300***	–0.0207***	0.00923	0.0598**
	(0.00346)	(0.00386)	(0.0203)	(0.0257)
Number of observations (pooled):	31,545	31,545	31,545	31,545

Notes: Statistical significance is indicated by p-values: : $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors are clustered by individual. Coefficient estimates of exogenous controls (X_{it} in Equation 1) have been omitted for brevity, but were present in the estimations of each model. Model 1 refers to the baseline model estimated via pooled OLS. Model 2 is the baseline via fixed effects. Model 3 is the corrected model via random effects, and Model 4 is the corrected model via fixed effects. The main results are reproduced from Table 3 for ease of comparison to the two alternate health indices. Section V provides details on the Bound et al. (1999) and health-limit health indices.

includes the more "broadly based" subjective self-reported health. This is an avenue for future research, as there are many possible methods for health measurement.

VI. Interpretation

This section explores the main findings in more detail. The first subsection relates changes in the health index to specific ailment conditions in order to better interpret the magnitude of the retirement effect. The next subsection investigates adaptions of the main model in order to observe the direct association of retirement with two factors—smoking and exercise—that may drive its beneficial influence on health. These analyses motivate some simple policy-related exercises at the end of the section, which demonstrate practical applications of the model.

A. Interpretation of Main Results

The previously provided interpretation for the corrected estimate of θ_2 was that an average length retirement spell (7.4 years, conditional on retirement having occurred two or more years before the latest survey) is associated with a health index that is 0.0194 points higher, holding all other observables fixed. Following the work of Angrist, Imbens, and Rubin (1996), this coefficient has an additional causal interpretation as the local average treatment effect (LATE) of retirement on health change. The calculated effect is attributable specifically to subpopulations who are affected by changes in the instruments. In other words, θ_2 represents the average effect of retirement on health changes for the subpopulation who responds to retirement preferences and age-based (62 and 65) incentives. An important caveat is that this empirical approach does not reveal anything about the retirement effect amongst the set of individuals who would always retire at a certain point regardless of those characteristics. Thus the LATE does not apply to individuals who experience involuntary retirement due to illness or injury.

A regression decomposition of the health index provides a framework for more concrete interpretation of the LATE. Table 8 presents a simple OLS regression of the health index on the nine ailments as well as self-reported health dummies. (It also presents regression results for the health indices used in the robustness checks.) Holding self-reported health constant, high blood pressure, diabetes, and heart problems carry the highest weight in the main health index. Cancer possesses a surprisingly low weight, perhaps implying that its effect is "washed out" by the self-reported health covariates. Applied to Table 8, the main retirement estimate $\hat{\theta}_2 = 0.0194$ indicates that the LATE of retirement on health change (over the average observed length of a long-term retirement spell, which is 7.4 years) is approximately equivalent to prevention of "one-quarter" of a doctor-diagnosed condition such as arthritis, or smaller fractions of more severe ailments. Alternatively, it implies that a 7.4 year long retirement spell should prevent arthritis for one in four individuals.

$\hat{\theta}_2$ represents a "pooled cumulative retirement effect" for the subpopulation that responds to the instrument set. Even within this subgroup, it is reasonable to conjecture that there may be strong heterogeneity in response to retirement, but due to the empirical limitations discussed in Section III, it is difficult to capture such effects.

Table 8
Health Index Decompositions

	Main Index	Health-limit Index	Bound et al. (1999) Index
Reference group: Self-reported health = 1 (excellent)			
Self-reported health = 2 (very good)	–0.0751***	–0.0210***	–0.00139***
	(0.000546)	(0.000976)	(0.000390)
Self-reported health = 3 (good)	–0.157***	–0.130***	–0.00246***
	(0.000682)	(0.00119)	(0.000465)
Self-reported health = 4 (fair)	–0.256***	–0.377***	–0.00470***
	(0.00104)	(0.00192)	(0.000652)
Self-reported health = 1 (poor)	–0.371***	–0.614***	–0.00575***
	(0.00205)	(0.00272)	(0.000990)
High blood pressure	–0.0906***	–0.00550***	–0.0896***
	(0.000560)	(0.000778)	(0.000402)
Diabetes	–0.0966***	–0.0279***	–0.136***
	(0.000785)	(0.00123)	(0.000528)
Cancer	–0.0251***	–0.0366***	–0.0754***
	(0.000882)	(0.00140)	(0.000635)
Lung problems	–0.0682***	–0.119***	–0.170***
	(0.00110)	(0.00204)	(0.000616)
Heart problems	–0.0919***	–0.0896***	–0.123***
	(0.000844)	(0.00126)	(0.000478)
Stroke	–0.0652***	–0.177***	–0.130***
	(0.00162)	(0.00340)	(0.00110)
Psychological problems	–0.0641***	–0.128***	–0.126***
	(0.000872)	(0.00166)	(0.000576)
Arthritis	–0.0769***	–0.111***	–0.0983***
	(0.000502)	(0.000863)	(0.000377)
Obese	–0.0470***	–0.0228***	–0.0530***
	(0.000527)	(0.000849)	(0.000375)
Constant	1.015***	1.030***	0.996***
	(0.000493)	(0.000936)	(0.000350)
Number of observations (pooled):	31,545	31,545	31,545

Notes: Statistical significance is indicated by p-values: : $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered by individual. The three columns present results from three regressions, each using a different health index. Section III provides details on calculation of the main health index. Section V provides details on the Bound et al. (1999) and health-limit health indices.

While this remains a question for future work, the next subsection explores possible channels through which the retirement effect may act.

B. Analysis of Health Behaviors

There is a small body of literature on the connection between leisure and health outcomes. Ruhm (2000) analyzed time series data regarding business cycles and mortality rates, determining that many types of fatalities, such as those caused by cardiovascular and liver disease, decrease when the economy deteriorates (aside from suicides). He also investigated microdata from the Behavioral Risk Factor Surveillance System (1987–95), concluding that "higher joblessness is associated with reduced smoking and obesity, increased physical activity, and improved diet." Ruhm's findings support that healthy habits may form an underlying mechanism through which the retirement-effect could act.

To this end, it is informative to examine individuals' smoking and exercise habits. Because retirees have more time to invest in their health, it may be easier for them to quit smoking or to be more physically active when not burdened by the work-week grind. Of the set of individuals who ever report smoking during the survey, 68.7 percent reported smoking during the interview that took place two to four years before their retirement (with 95 percent confidence interval of [65.2 percent, 72.2 percent]). For the same set of individuals, only 56.3 percent reported smoking two to four years after their retirement (with 95 percent confidence interval of [52.7 percent, 59.9 percent]). The corresponding statistics for those reporting at least 30 minutes of vigorous exercise three or more days per week are: 47.9 percent at two to four years before retirement (95 percent C.I. of [45.3 percent, 50.4 percent]) and 51.6 percent at two to four years after retirement (95 percent C.I. of [48.9 percent, 54.3 percent]).[31] Thus smoking incidence declines postretirement (statistically significantly), while exercise levels appear to rise. On the other hand, the data reveal that these two health behaviors improved on average for the entire sample throughout the panel, likely reflecting changing attitudes toward smoking and exercise habits during the 1990s and 2000s. Thus it is unclear whether retirement or some other variable is driving these stylized facts.

Using regression analysis with smoking and exercise on the lefthand-side, observable variation can be conditioned upon (for example, a time trend could explain changing attitudes toward smoking and exercise), but endogeneity problems, comparable to those in the baseline econometric model, may remain. For instance, a retirement-causing health decline may induce a lower level of physical activity, or it may elicit a doctor's order to quit smoking. In the case of exercise, such a scenario would work against the direction of the retirement effect implied by the statistics above, biasing OLS estimates of the retirement's impact on exercise levels downward. The "doctor's orders" effect would have the opposite impact. In general, it is not possible to heuristically determine the sign of the bias.[32]

31. The exercise statistics omit observations from the 2004 wave and on, due to phrasing changes in the survey questionnaire which led to inconsistencies in the data.

32. One can imagine converse effects. For instance, the stress of a severe illness could compel an individual to resume smoking when he or she had previously quit. Or, a retirement-causing disability might necessitate physical therapy treatments classified as "vigorous activity."

As a starting point, Table 9 presents results from the following models estimated as both random and fixed effects logit specifications:

(9) $VigAct_{it} = F(\alpha t + \beta X_{it} + \theta_1 RS_{it} + \theta_2 RL_{it} + f_i + \varepsilon_{it})$

(10) $Smokes_{it} = F(\alpha t + \beta X_{it} + \theta_1 RS_{it} + \theta_2 RL_{it} + f_i + \varepsilon_{it})$

In these models, parameters and variables are defined as in Equations 1 and 2 in Section IV. $F(\cdot)$ is the logistic function, a time trend, t, is now an explanatory variable, and ε_{it} encompasses all unobserved time variant information. The table contains estimates for both random and fixed effects logit models. The sample sizes are reduced because vigorous activity models were estimated on pre-2004 observations (the survey question was altered in later waves), and smoking models were estimated only on respondents who smoked during at least one survey wave. Furthermore, the fixed effects samples are even smaller, because the estimations utilize variation only from individuals who switch their (dependent variable) health behavior during the panel.

The fixed effects logit models yield reasonable estimates of the covariates. Wealthier individuals tend to exercise more, and married respondents smoke less than unmarried ones. The time trend is correlated with higher exercise levels and lower smoking levels, although it is not significant in the smoking models. In fixed effects models, both short- and long-term retirement estimates possess the signs implied by the stylized facts: Retirement is associated with more exercise and less smoking, and the effect is most significant for long-term retirement. For vigorous activity models, coefficient estimates switch sign, going from RE to FE. It is important to note that these results are merely correlational; while it is possible to speculate on the sign of the biases in the retirement estimates, it is not clear that the instrumental variables strategy of Section V would properly correct endogeneity in these models. In general, this discussion explores some underlying features of the retirement-health mechanism, while deeper study into the cause and effect of health behaviors is a possible direction for future research. Further investigation may also reveal that individuals alter other health behaviors following retirement, such as preventive medical care, drinking, and nutrition.

C. Application: Social Security

Much debate has centered on questions regarding entitlement systems in the United States. For example, due to changes in demographics and the labor market, the long-run trend in Social Security benefit payouts will begin to exceed pay-ins. Under the status quo, the Social Security trust fund will shrink and vanish over the next few decades, at which time it will no longer be possible to fully fund the system. Researchers, politicians, and government agencies have posed many different reform plans, including tax increases and benefit decreases. This paper's main results should be considered when assessing the impact of such policy recommendations.

Schemes to increase the FICA tax would likely affect retirement behavior. A percentage-wise tax increase may compel workers to retire later in order to increase their lifetime earnings, or conversely it may provide a disincentive for work. Removal of the cap on taxable earnings (in 2012, annual earnings above \$110,100 were not

Table 9
Health Behavior Models

	Vigorous Activity		Smoking	
	Logit (RE)	Logit (FE)	Logit (RE)	Logit (FE)
Female	–0.494***		0.0000834	
	(0.0678)		(0.189)	
t	0.209***	0.923***	–0.0569	–0.156
	(0.0222)	(0.301)	(0.0347)	(0.406)
Age	0.636***	0.822***	–0.155	0.0126
	(0.137)	(0.238)	(0.149)	(0.273)
Age^2	–0.00520***	–0.00974***	–0.000319	–0.00126
	(0.00112)	(0.00143)	(0.00115)	(0.00127)
Black	–0.0799		–0.435*	
	(0.0993)		(0.257)	
Hispanic	–0.0961		–0.464	
	(0.137)		(0.390)	
Married	0.0594	–0.0114	–0.649***	–0.554**
	(0.0741)	(0.176)	(0.153)	(0.221)
Log(assets)	0.113***	0.0285	–0.0415	–0.0157
	(0.0161)	(0.0254)	(0.0254)	(0.0318)
Log(debt)	0.106***	0.0429	–0.0339	–0.0172
	(0.0249)	(0.0325)	(0.0337)	(0.0388)

Some high school	0.00517		0.323	
	(0.156)		(0.400)	
High school diploma	0.00107		–0.195	
	(0.140)		(0.367)	
Some college	0.0925		–0.356	
	(0.148)		(0.389)	
Bachelor's degree	–0.0771		–0.731	
	(0.165)		(0.469)	
Graduate degree	0.0101		–0.919**	
	(0.159)		(0.468)	
RL_{it}: Long-term retirement	–0.100	0.227*	–0.521***	–0.393**
	(0.0790)	(0.127)	(0.150)	(0.181)
RS_{it}: Short-term retirement	–0.125	0.0484	–0.343**	–0.273*
	(0.0772)	(0.0949)	(0.144)	(0.155)
Number of observations (pooled):	13,872	7,271	7,868	3,909

Notes: Statistical significance is indicated by p-values: : $*p < 0.10$, $**p < 0.05$, $***p < 0.01$. The first two columns contain random and fixed effects logit estimates for Equation 9, respectively. The third and fourth columns contain the corresponding estimates for Equation 10. Coefficient estimates of the constant term have been omitted for brevity, but were present in the estimations of all four models. The reference category for education dummy variables is "zero years of high school."

subject to the Social Security tax) may have similar effects for higher income workers. Alternatively, an increase in the age of eligibility to receive benefits would compel all workers to retire later. A number of studies have forecast the impact of such possible legislation. Coile and Gruber (2007) considered two changes: the effect of an instantaneous increase in the normal retirement age (NRA[33]) from 65 to 67 (the NRA was 65 for all individuals in their sample although it has undergone a phased-in increase to 67 for younger Americans), and an increase in the delayed retirement credit (DRC) by three percentage points.[34] For the first reform measure, they predicted a 2 percent increase in the labor force participation rate for individuals between ages 65 and 67 (and more moderate increases for individuals near those ages). In the latter reform, they forecast a 4 percent jump in the labor force participation rate. Samwick (1998) estimated similar effects on the retirement rate. He predicted a one percentage point reduction in the retirement rate due to either a 20 percent decrease in the Social Security benefit or instantaneous implementation of the 1983 amendments to Social Security (which have been gradually phased in).[35]

The main results of Section V provide a method to measure the second-order effects of such Social Security changes on health aggregates.[36] For example, researchers may be interested in predicting how these types of reforms might affect healthcare spending associated with high blood pressure. Suppose a particular reform option increased the retirement rate by one percentage point over two years. According to Table 3, such an event would yield higher average health by one percent of the estimate of the coefficient of RL_{it}, which corresponds to an average increase of 0.000194 in the health index across the entire sample. Table 8 shows that a diagnosis of high blood pressure is, on average, associated with a decrease in the index by 0.0906. Consequently, the reform is expected to decrease the incidence of high blood pressure by 0.214 percent (= 0.000194/0.0906) in the retirement-aged labor force. These types of calculations could be performed for more ailment conditions, and they illustrate simple ways to explore some deeper effects of possible reforms. Moreover, they suggest an avenue for future research.

VII. Conclusion

This paper endeavors to resolve the debate on the influence of retirement on health changes. The main challenge is to overcome the two-way causal connection between health status and retirement. To examine the effect, a small body of previous literature has utilized fixed effects and instrumental variables estimation with

33. NRA is defined as the age at which retirement benefits are equal to the "primary insurance amount," which is the standard benefit package. Details on its size are available at www.socialsecurity.gov.
34. From 5 percent to 8 percent for individuals in their sample. The DRC is a percentage increase in monthly benefits that increases as a worker continues employment past the NRA but before age 70.
35. The 1983 amendments consisted of three components: (1) An increase in the normal retirement age from 65 to 67; (2) a larger reduction in benefits for opting into early retirement at age 62; (3) a smaller delayed retirement credit for postponing retirement to age 70.
36. "First-order" effects refer to the impact of entitlement reform on retirement behavior. Second-order effects are the resultant health effects due to those changes in retirement behavior.

limited success. Fixed effects studies, such as Dave, Rashad, and Spasojevic (2008), did not fully treat the simultaneous effects problem, so they yielded biased estimates. Other research performed by Neuman (2008) and Coe and Zamarro (2011) adopted discontinuous retirement incentives due to private and public benefit plans as instrumental variables. They measured a small positive effect on health but struggled to obtain significant results, while using a variety of health measures. This paper employs a stronger instrument—individuals' retirement expectations—to estimate the impact of retirement on a robust health index that incorporates both objective and subjective health characteristics. The instrument set exploits useful information that can be found in subjective expectations variables, increasingly popular additions to survey questionnaires. The main results reveal that IV estimates switch sign when compared to OLS estimates. In summary, retirement exerts a beneficial influence on health changes that is approximately equivalent to prevention of "one-quarter" of a doctor-diagnosed condition, such as diabetes, or various fractions of other ailments. Evidence also suggests that retirement acts on health through beneficial behaviors. In particular, individuals who retire apply more effort to quit smoking.

A direct and measurable impact of retirement on health suggests many future directions for research. Through retirement, health changes are key byproducts of potential reforms to expansive programs such as Social Security and Medicare. More specific examination of the data may yield an even deeper understanding of these effects. Exploration of particular ailments, diseases, mortality rates, and medical care spending levels may reveal further causal interactions with retirement. The question of how retirees spend their newfound leisure time is also compelling. Finally, there are new questions regarding retirement motives: Given a clearer understanding of retirement's influence on health, previous studies regarding the influence of health and economic factors on retirement should be revisited. Together, these new insights may shape policies and help generate economic stability and prosperity amidst large-scale demographic changes.

References

Anderson, Kathryn, and Richard Burkhauser. 1985. "The Retirement-Health Nexus: A New Measure of an Old Puzzle." *Journal of Human Resources* 20(3):315–30.

Angrist, Joshua, Guido Imbens, and Donald Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–55.

Baker, Michael, Mark Stabile, and Catherine Deri. 2004. "What Do Self-Reported, Objective, Measures of Health Measure?" *Journal of Human Resources* 39(4):1067–93.

Bazzoli, Gloria. 1985. "The Early Retirement Decision: New Empirical Evidence on the Influence of Health." *Journal of Human Resources* 20(2):214–34.

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

Bound, John, and Timothy Waidmann. 2007. "Estimating the Health Effects of Retirement." Michigan Retirement Research Center Research Paper No. UM WP 2007-168. Ann Arbor: University of Michigan.

Buchmueller, Thomas, and Robert Valletta. 1999. "The Effect of Health Insurance on Married Female Labor Supply." *Journal of Human Resources* 34(1):42–70.
Charles, Kerwin Kofi. 2004. "Is Retirement Depressing? Labor Force Inactivity and Psychological Well Being in Later Life." *Research in Labor Economics* 23:269–99.
Coe, Norma, and Gema Zamarro. 2011. "Retirement Effects on Health in Europe." *Journal of Health Economics* 30(1):77–86.
Coile, Courtney, and Jonathan Gruber. 2007. "Future Social Security Entitlements and the Retirement Decision." *Review of Economics and Statistics* 89(2):234–46.
Dave, Dhaval, Inas Rashad, and Jasmina Spasojevic. 2008. "The Effects of Retirement on Physical and Mental Health Outcomes." *Southern Economic Journal* 75(2):497–523.
Dwyer, Debra, and Olivia Mitchell. 1999. "Health Problems as Determinants of Retirement: Are Self-Rated Measures Endogenous?" *Journal of Health Economics* 18:173–93.
Gustman, Alan, and Thomas Steinmeier. 1994. "Employer Provided Health Insurance and Retirement Behavior." *Industrial and Labor Relations Review* 48(1):124–40.
———. 2000. "Retirement Outcomes in the Health and Retirement Study." *Social Security Bulletin* 63(4):57–71.
Jurges, Hendrik. 2007. "True Health vs. Response Styles: Exploring Cross-Country Differences in Self-Reported Health." *Health Economics* 16(2):163–78.
Lange, Fabian, and Douglas McKee. 2011. "The Evolution of Latent Health over the Life Course." Unpublished.
McDowell, Ian. 2006. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 3rd edition. New York: Oxford University Press.
McHorney, Colleen, and Allan Cohen. 2000. "Equating Health Status Measures with Item Response Theory: Illustrations with Functional Status Items." *Medical Care* 38(9 Supplement II):II43–59.
McGarry, Kathleen. 2004. "Health and Retirement: Do Changes in Health Affect Retirement Expectations?" *Journal of Human Resources* 39(3):624–48.
Neuman, Kevin. 2008. "Quit Your Job and Get Healthier? The Effect of Retirement on Health." *Journal of Labor Research* 29(2):177–201.
Olin, Gary, and Jeffrey Rhoades. 2006. "The Five Most Costly Medical Conditions, 1997 and 2002: Estimates for the U.S. Civilian Noninstitutionalized Population." Statistical Brief #80. Rockville: Agency for Healthcare Research and Quality.
Olson, Craig. 1998. "A Comparison of Parametric and Semiparametric Estimates of the Effect of Spousal Health Insurance Coverage on Weekly Hours Worked by Wives." *Journal of Applied Econometrics* 13(5):543–65.
RAND HRS Data, Version L. 2011. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica.
Rohwedder, Susann, and Robert Willis. 2010. "Mental Retirement." *Journal of Economic Perspectives* 24(1):119–38.
Ruhm, Christopher. 2000. "Are Recessions Good for Your Health?" *Quarterly Journal of Economics* 115(2):617–50.
Rust, John, and Christopher Phelan. 1997. "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets." *Econometrica* 65(4):781–831.
Samwick, Andrew. 1998. "New Evidence on Pensions, Social Security, and the Timing of Retirement." *Journal of Public Economics* 70:207–36.
Madrian, Brigitte, Olivia Mitchell, and Beth Soldo, eds. 2007. *Redefining Retirement: How Will Boomer Fair?* Oxford: Oxford University Press.
Stanton, Mark. 2006. "The High Concentration of U.S. Health Care Expenditures." *Research in Action Issue 19*. AHRQ Publication No. 06-0060. Rockville: Agency for Healthcare Research and Quality.

Ware, John, Mark Kosinski, Martha Bayliss, Colleen McHorney, William Rogers, and Anastasia Raczek. 1995. "Comparison of Methods for the Scoring and Statistical Analysis of SF-36 Health Profile and Summary Measures: Summary of Results from the Medical Outcomes Study." *Medical Care* 33(4):264–79.

Zhan, Yujie, Mo Wang, Songqi Liu, and Kenneth Schultz. 2009. "Bridge Employment and Retirees' Health: A Longitudinal Investigation." *Journal of Occupational Health Psychology* 14(4):374–89.