Health, Health Insurance, and Retirement: A Survey

Eric French  John Bailey Jones*

January 25, 2017

Abstract

The degree to which retirement decisions are driven by health is a key concern for both academics and policymakers. In this paper we survey the economic literature on the health-retirement link in developed countries. We describe the mechanisms through which health affects labor supply and discuss how they interact with public pensions and public health insurance. The historical evidence suggests that health is not the primary source of variation in retirement across countries and over time. Furthermore, declining health with age can only explain a small share of the decline in employment near retirement age. Health considerations nonetheless play an important role, especially in explaining cross-sectional variation in employment and other outcomes within countries. We review the mechanisms through which health affects retirement and discuss recent empirical analyses.

Keywords: Disability; Elderly; Health; Retirement; Social Security

*Eric French: UCL, CEPR, and IFS, eric.french.econ@gmail.com. John Bailey Jones: Federal Reserve Bank of Richmond and University at Albany, SUNY; jbjones@albany.edu. For helpful comments, we are grateful to Karen Kopecky and seminar participants at the Federal Reserve Bank of Richmond. The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the IFS. This survey has been prepared for the Annual Review of Economics.
1 Introduction

Old age is characterized by declining health and reduced labor supply. How are these events related? Although everyone eventually becomes too sick to work, people may retire well before then. The health-retirement link is particularly germane to the reform of public pension programs. Responding to fiscal pressures, many countries have restructured their pension programs to encourage later retirement (French and Jones, 2012). Such reforms are unlikely to be successful, however, if older individuals are too unhealthy to significantly extend their careers. On the other hand, if individuals are retiring in good health, they may be receptive to incentives for continued work.

In this paper we review the evidence for developed countries on the health-retirement link and its implications for public policy. Extending our work in earlier, more general reviews (French and Jones 2012, Blundell et al. 2016a), we examine the relationship between health and retirement in some detail. The literature contains a variety of methodologies, ranging from broad historical reviews to fully articulated dynamic models. This allows us to assess economic mechanisms as well as document empirical relationships.

Over the past century, there has been a pronounced decline in the labor supply of older men. Given that life expectancy (even at older ages) increased over the same period, the decline in labor supply is clearly not due to health.¹ Much of the decline is due instead to increases in retirement income, including but not limited to significant increases in public pensions. Even in the past two decades, when this decline has stopped and in many countries reversed, employment appears to have in general grown more slowly than “work capacity” (Coile et al., forthcoming).

The effects of health on retirement are more apparent in cross-sectional relationships. All else equal, people in worse health are less likely to work. Several mechanisms are possible. Illness can make work extremely unpleasant. Bad health can also reduce worker

¹As discussed in section 2 below, these increases in expected lifespans have also been increases in expected healthy lifespans.
productivity. People in sufficiently bad health may receive disability benefits, and people receiving disability benefits are usually prohibited from working. Finally, with shorter expected lifespans, individuals in bad health may not need to work as long accumulating financial and pension wealth for their retirements.

Health may also affect labor supply through medical expenses. Because many US workers only receive health insurance while they continue to work, expensive medical conditions may lead them to delay retirement. Alternatively, uninsured workers may leave their jobs in order to qualify for health care provided through disability or means-tested social insurance.

The rest of the paper is organized as follows. Section 2 documents the patterns of health and labor supply for older workers over the past century. In section 3 we survey recent evidence on how cross-sectional differences in health affect the labor supply and wages of older workers. In section 4 we use a structural model to illustrate how health affects retirement. In section 5 we discuss findings from recent structural studies. In Section 6 we conclude and offer some suggestions for future work.

2 Historical trends

The health of older individuals has improved dramatically over the past 90 years. Figure 1 shows that the life expectancy of 65-year-old men in the US and the UK has risen from about 11 years in 1920 to over 18 years today. Longer lifespans do not automatically imply longer potential working lives, as individuals may be spending more of their lives incapacitated by illness. A large body of work shows, however, that people are carrying their good health into older ages as well. Table 1, taken from Costa (2002), shows functional limitation rates for veterans of the Union Army (from the US Civil War) and for older men in the National Health and Nutrition Examination Surveys (NHANES). The NHANES cohort, born roughly 85 years after the Union Army veterans, has considerably
lower limitation rates. By way of example, 28.5% of the Union Army veterans aged 50-64 had difficulty walking, nearly three times the fraction in the NHANES.\(^2\) Recent studies utilizing detailed modern datasets show that healthy lifespans continue to increase. Chernew et al. (2016) find that between 1992 and 2008, healthy lifespans in the US grew by 1.8 years. Looking across countries and time, Milligan and Wise (2012, Figure I.21) show that among men aged 60-64, recent decreases in mortality rates correlate closely with decreased self-reports of fair or poor health. Results from the Global Burden of Disease Study (Vos et al., 2015) suggest that much of the worldwide increase in life expectancy is an increase in healthy life expectancy.\(^3\) The unresolved question appears not to be whether unhealthy lifespans are growing more slowly than total lifespans, but whether they are growing at all (Chernew et al., 2016).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ages 50-64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty bending</td>
<td>44.4</td>
<td>7.5</td>
</tr>
<tr>
<td>Difficulty walking</td>
<td>28.5</td>
<td>10.4</td>
</tr>
<tr>
<td>Paralysis</td>
<td>4.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Blindness in at least one eye</td>
<td>3.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Deafness in at least one ear</td>
<td>3.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Ages 60-74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty bending</td>
<td>53.8</td>
<td>16.1</td>
</tr>
<tr>
<td>Difficulty walking</td>
<td>36.6</td>
<td>13.8</td>
</tr>
<tr>
<td>Paralysis</td>
<td>6.1</td>
<td>2.7</td>
</tr>
<tr>
<td>Blindness in at least one eye</td>
<td>4.5</td>
<td>3.1</td>
</tr>
<tr>
<td>Deafness in at least one ear</td>
<td>4.1</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Table 1: Functional limitation rates for older men in percent. Source: (Costa, 2002, Table 3).

\(^2\)Costa (2002) constructs three sets of limitation rates from the Union Army data, to control for various measurement issues. Table 1 displays the benchmark measures, but the other measures are similar. She further argues “that several tests indicate that this sample is representative of the general population in terms of wealth and circa 1900 in terms of mortality experience.” Costa (2002) also calculates limitation rates for a few conditions from the 1994 National Health Interview Survey (NHIS), which resemble their NHANES counterparts.

\(^3\)Also see Costa (1998), Manton et al. (2008) and Cutler et al. (2014).
These improvements in health, however, have not lead to increased work. Figure 2 shows long-term employment trends for men aged 65 and older in the US and the UK. Between 1880 and 1980 the employment rate for these men fell from three-quarters to less than one-fifth. France and Germany experienced similar trends (Costa, 1998). In her historical review, Costa (1998) concludes that the labor supply of the elderly has fallen for two reasons. The first is that decreases in the cost of recreational goods, along with an increased capacity for independent living, have made retirement more attractive. The second is that retirement income has grown. Older individuals have become increasingly able to support themselves through personal savings, private pensions, and public pensions.

A number of theoretical papers analyze how increased longevity affects retirement. Bloom et al. (2014) provide a particularly intuitive treatment. They point out that a key factor is the extent to which longer lifespans lead to larger potential lifetime earnings. As long as growth in potential earnings is at least “proportional” to the growth in life

Source: UK data from the Office for National Statistics. US data from the Human Mortality Database.
expectancy, and leisure is a normal good, workers should respond to longer lifespans by lengthening their retirements. Bloom et al. (2014) show that retirement ages should in fact rise less than proportionally with expected lifespans, as compounding interest allows workers to retire at relatively younger ages. Nonetheless, Bloom et al. (2014), along with most other analyses, conclude that increased longevity should cause retirement ages to rise rather than fall.\footnote{Other papers include Hammermesh (1984), Chang (1991), Kalemli-Ozcan and Weil (2010) and d’Albis et al. (2012).}

Several recent papers assess Costa’s hypotheses with quantitative models. All of these analyses consider the sustained increase in wages that has occurred over the past century. If the income effects of these wage increases dominate the substitution effects, the fraction of life spent in retirement should rise. Calibrating their model to the wage increases and
mortality declines observed in the US, Bloom et al. (2014) find that it reproduces the decline in retirement ages observed over the past 50 years. Restuccia and Vandenbroucke (2013) conclude that almost all of the decrease in working hours between the 1870 and 1970 birth cohorts can be attributed to higher wages. Kopecky (2011) considers the observed decline in the price of leisure goods, in addition to the increase in wages, in a model where time and leisure goods are combined to provide leisure services. She finds that together these changes can explain most of the observed increase in retirement. While higher wages are the dominant factor, cheaper leisure goods are also important.

Much attention has been given to another secular trend, the expansion of public pensions. When they were introduced in 1889, the eligibility age for public pensions in Germany was 70. In 1916 the age was lowered to 65. Today, workers in Germany with sufficiently long earnings histories can retire at age 63, and unemployment and disability insurance allow for even earlier retirement (Börsch-Supan and Jürges 2012; Kemptner 2016). In the UK the eligibility age fell from 70 in 1909 to 65 in 1925. While the initial eligibility age for the US Social Security program was 65, in 1961 workers were allowed to retire at age 62, in exchange for lower annual benefits. In addition to allowing older individuals to afford longer retirements – a standard wealth effect – many public pension programs generate substitution effects that encourage retirement. For example, Social Security recipients traditionally faced the Earnings Test, where earnings above a certain threshold led to reduced benefits. Moreover, Social Security recipients face higher effective income tax rates, because the portion of Social Security benefits subject to income taxation is increasing in total income (Jones and Li, 2016). Applying a comprehensive “Fiscal Analyzer” to data from the Survey of Consumer Finances, Auerbach et al. (2016) find that many older workers face marginal net tax rates of 50% or higher.

---

5 Even though benefits lost through the Earnings Test were credited to future benefits, the credits were generally less than actuarially fair, so that the Earnings Test was on net an earnings tax. The empirical consensus is that the partial elimination of the Earnings Test in 2000 increased labor supply (Engelhardt and Kumar, 2014).
More recent evidence from an expanded set of countries points to similar conclusions. Blundell et al. (2016a) show employment rates over the past 50 years for men and women in a number of developed countries. In most countries, the trend of employment has been U-shaped for those aged 60-64, falling until the 1990s and rising afterward. In contrast, health improved throughout the entire period. Coile et al. (forthcoming) document how the relationship between health and work has changed over time. In each of 12 countries, they find employment and mortality – a proxy for health – at each age in a particular calendar year, allowing them to express employment as a function of mortality. They find that employment at any given value of mortality fell much more between 1977 and 1995 than between 1995 and 2010 (Coile et al., forthcoming, Figure 3 and Table 3). In fact, while employment at older ages fell between 1977 and 1995, between 1995 and 2010 employment at older ages rose. These changes in employment generally coincided with changes in public pension systems. Until the 1990s public pensions evolved in ways that discouraged work at older ages, while more recent reforms have encouraged work (French and Jones 2012; Blundell et al. 2016a). Using a structural general equilibrium model, Ferreira and dos Santos (2013) study the increase in retirement experienced in the US between 1950 and 2010. They conclude that most of the increase was due to enhancements to Social Security and the introduction of Medicare health insurance for the elderly.

Cross-country analyses corroborate the importance of pensions. Proceeding similarly to Coile et al. (forthcoming), Milligan and Wise (2012) calculate employment as a function of mortality for 2005. Comparing seven developed countries, they show that the functions differ greatly (Figure I.9), suggesting strongly that health cannot be the sole determinant of retirement. Gruber and Wise, eds (2004) show that in the 11 countries they study labor force exits are concentrated around legislated early and normal pension retirement ages. Duval (2003) calculates the decrease in lifetime benefits and additional payroll taxes that workers incur when they delay claiming their public pensions. He finds that employment rates at older ages tend to be lower in countries where these implicit taxes are high. Using
dynamic structural models, Erosa et al. (2012), Wallenius (2013), Alonso-Ortiz (2014) and Laun and Wallenius (forthcoming) all conclude that much of the variation in retirement across developed countries can be explained by differences in their public pension systems.

Because pensions likely reflect societal preferences for retirement (e.g., Bloom et al. 2014), identifying their independent contribution to retirement trends is difficult. But in any event, the long-term trend in retirement has not followed the long-term trend in health. Although bad health undoubtedly continues to force many older people out of the labor market, many other retirees could still work if they so chose. Moreover, the variation in retirement across developed countries appears to be due mostly to institutional differences.

3 Cross-sectional relationships

3.1 Health measurement

Even if health has not been the key driver of longer-term retirement trends, it may well have significant predictive power at the household or individual level. Any analysis of these micro-level relationships, however, is complicated by difficulties in measuring health itself. Older surveys containing employment data often contained little health-related data. Fortunately, many large surveys – such as the Health and Retirement Study (HRS), the English Longitudinal Study of Ageing (ELSA) and their international counterparts – now contain numerous subjective and objective health measures. But even these measures have their limitations.

Table 2 presents summary statistics for a number of health measures in the ELSA. The first set of measures contains objective measures of specific health conditions. While these measures are accurate – there is little ambiguity as to whether an individual has

---

6 More detailed discussions of these issues can be found in Currie and Madrian (1999), O'Donnell et al. (2015) and Blundell et al. (2016a).
received cancer treatments – they suffer from several conceptual difficulties. First, objective measures are specific rather than comprehensive. Important medical conditions may not be reported. Second, many objective measures are binary variables that do not capture the severity of the health problems. Third, the link between any specific objective health condition and labor supply can depend as much on the individual’s occupation as on the condition’s severity. Arthritis that would sideline a roofer may not hinder an office worker at all. As a result, health conditions alone are noisy signals of “work capacity” (Currie and Madrian, 1999). These problems are all variants of classical measurement error, which likely bias the estimated effect of health towards zero.

Included in our list of “objective measures” are self-reports of one’s ability to perform Activities of Daily Living (ADLs) such as walking, getting dressed, etc. These measures are not clearly objective, as they are in no way based upon a doctor’s diagnosis, but they are linked to the ability to perform specific tasks. Related measures include Instrumental Activities of Daily Living, such as cooking or managing money, and cognitive measures, such as remembering words. While these variables are not listed in Table 2, they can be found in the HRS, ELSA, and related surveys.

The second set of measures shown in Table 2 are subjective measures. These include self-assessed health status – reports of whether the respondent’s health is “excellent, very good, good, fair or poor” – and self-reports of whether the individual is limited in their ability to do work or other activities. Subjective measures have the benefit of being simple measures that incorporate all health conditions. In addition, subjective measures are better directed to economic concepts such as work capacity. Recalling the example above, if a particular arthritic condition impedes a roofer but not an office worker, the roofer should report a work limitation while the office worker should not. However, subjective measures are always subject to reporting error, and they are often relatively crude measures, taking on only a handful of values. If unbiased, the measurement error in subjective measures will result in an attenuation effect parallel to that associated with
objective measures. On the other hand, health measures may also suffer from “justification bias,” where those who are not working inaccurately claim to be unhealthy in order to justify their work status (see, e.g., Butler et al. (1987)). This would likely lead researchers to overstate the effect of health on labor supply.

Stern (1989) suggests using objective health measures to instrument for more subjective measures. Bound (1991) shows that this procedure produces estimates that are close to those found by simply using the subjective measures (although the procedure can affect the estimated value of financial incentives). This suggests that for the subjective health measures the effects of measurement error and justification bias roughly offset. Blundell et al. (2016b) come to similar conclusions. Dwyer and Mitchell (1999) and McGarry (2004) circumvent concerns about justification bias by examining the relationship between health and expected, rather than realized, retirement. Since in this approach retirement has not been realized, there is little need to justify retirement status by reporting bad health. They continue to find strong links between subjective health measures and retirement. Benitez-Silva et al. (2004) find that they cannot reject the hypothesis that self reported disability is an unbiased measure of true disability, although Kreider and Pepper (2007) find that nonworkers tend to overreport disability rates.
Table 2: ELSA data - description of health variables for selected sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>N</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective health measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>Received cancer treatment in past 2 years</td>
<td>41,361</td>
<td>0</td>
<td>1</td>
<td>0.02</td>
<td>0.15</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Taking medication for diabetes</td>
<td>41,356</td>
<td>0</td>
<td>1</td>
<td>0.05</td>
<td>0.22</td>
</tr>
<tr>
<td>Sight</td>
<td>Reported poor eyesight</td>
<td>41,358</td>
<td>0</td>
<td>1</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td>Hearing</td>
<td>Reported poor hearing</td>
<td>41,360</td>
<td>0</td>
<td>1</td>
<td>0.03</td>
<td>0.17</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>Taking medication for high blood pressure</td>
<td>41,389</td>
<td>0</td>
<td>1</td>
<td>0.24</td>
<td>0.42</td>
</tr>
<tr>
<td>Arthritis</td>
<td>Reported arthritis this wave</td>
<td>41,154</td>
<td>0</td>
<td>1</td>
<td>0.28</td>
<td>0.45</td>
</tr>
<tr>
<td>Psychiatric</td>
<td>Reported psychiatric problem this wave</td>
<td>41,391</td>
<td>0</td>
<td>1</td>
<td>0.07</td>
<td>0.25</td>
</tr>
<tr>
<td>Difficulty Walking One Block</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.09</td>
<td>0.28</td>
</tr>
<tr>
<td>Difficulty Sitting for Two Hours</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.13</td>
<td>0.34</td>
</tr>
<tr>
<td>Difficulty Getting Up from a Chair</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.21</td>
<td>0.41</td>
</tr>
<tr>
<td>Difficulty Climbing Several Flights of Stairs</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.28</td>
<td>0.45</td>
</tr>
<tr>
<td>Difficulty Climbing One Flight of Stairs</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.10</td>
<td>0.30</td>
</tr>
<tr>
<td>Difficulty Stooping, Kneeling, or Crouching</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.30</td>
<td>0.46</td>
</tr>
<tr>
<td>Difficulty Lifting or Carrying 10 pounds</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.18</td>
<td>0.38</td>
</tr>
<tr>
<td>Difficulty Picking Up a Dime</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.04</td>
<td>0.20</td>
</tr>
<tr>
<td>Difficulty Extending Arms</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.09</td>
<td>0.29</td>
</tr>
<tr>
<td>Difficulty Pushing or Pulling Large Object</td>
<td>Mobility: Does not (0), does (1)</td>
<td>41,297</td>
<td>0</td>
<td>1</td>
<td>0.13</td>
<td>0.34</td>
</tr>
<tr>
<td>Difficulty Walking across Room</td>
<td>ADL: Does not (0), does (1)</td>
<td>41,299</td>
<td>0</td>
<td>1</td>
<td>0.02</td>
<td>0.15</td>
</tr>
<tr>
<td>Difficulty Getting Dressed</td>
<td>ADL: Does not (0), does (1)</td>
<td>41,299</td>
<td>0</td>
<td>1</td>
<td>0.10</td>
<td>0.30</td>
</tr>
<tr>
<td>Difficulty Bathing or Showering</td>
<td>ADL: Does not (0), does (1)</td>
<td>41,299</td>
<td>0</td>
<td>1</td>
<td>0.07</td>
<td>0.26</td>
</tr>
<tr>
<td>Difficulty Eating</td>
<td>ADL: Does not (0), does (1)</td>
<td>41,299</td>
<td>0</td>
<td>1</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>Difficulty Getting In or Out of Bed</td>
<td>ADL: Does not (0), does (1)</td>
<td>41,299</td>
<td>0</td>
<td>1</td>
<td>0.05</td>
<td>0.23</td>
</tr>
<tr>
<td>Difficulty Using the Toilet</td>
<td>ADL: Does not (0), does (1)</td>
<td>41,299</td>
<td>0</td>
<td>1</td>
<td>0.03</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>Subjective health measures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health limits activities</td>
<td>Does not (0), does (1)</td>
<td>39,421</td>
<td>0</td>
<td>1</td>
<td>0.53</td>
<td>0.50</td>
</tr>
<tr>
<td>General health</td>
<td>Excellent (1), very good (2), good (3), fair (4), poor (5)</td>
<td>36,231</td>
<td>1</td>
<td>5</td>
<td>2.59</td>
<td>1.11</td>
</tr>
<tr>
<td>Health limits work</td>
<td>Does not (0), does (1)</td>
<td>33,341</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>0.43</td>
</tr>
</tbody>
</table>
An important issue is whether the effects of health on economic outcomes are best studied using cross-sectional methods or panel data methods. Cross-sectional estimates may overstate the effect of health on work and wages because of reverse causality. This reverse causality could come from a number of sources. First, it could be that higher incomes cause better health. Those with higher income may be able to purchase better nutrition and health care, for example, or face less stress. Indeed, epidemiologists attribute most of the correlation between health and income to income’s effect on health, rather than vice-versa: see Brunner (2016) for a recent review discussing how low social status affects the risk of cardiovascular disease and other stress-related health problems. Second, it could be due to other factors that lead to both good health and high earnings. It may be the case that those who are more patient tend to both work more and invest more in their health. Likewise, high-income parents may invest more in both the health and the education of their children, leading health and income to be positively correlated at older ages. Case et al. (2002) show that not only is children’s health positively related to household income, but the relationship between household income and children’s health becomes more pronounced as children age. Part of the relationship can be explained by the arrival and impact of chronic conditions.

One approach to this problem is to use panel data approaches, such as fixed-effects, first differencing, or controlling for initial health and other conditions. These panel data estimators exploit contemporaneous changes in health and employment. It seems plausible that most of these changes are the result of health shocks affecting work and wages, rather than sudden declines in employment affecting health.

Panel data approaches lead to much smaller estimates of the effects of health on work and wages than do cross-sectional estimators. When estimating the effect of health on employment, Blundell et al. (2016b) show that OLS estimates of the effect of health on employment are close to 10 times as large as first difference estimates. This suggests that cross-sectional estimates overstate the effect of health on work and wages.
However, there are reasons to believe that first difference estimates understate the true causal effects of health. Perhaps most important is that the measurement error in health is likely to be large, for the reasons given above, leading first difference estimators to suffer from attenuation bias. Related to the issue of measurement error is the issue that some health shocks are relatively transitory, whereas other shocks are more permanent. Blundell et al. (2016b) show that the employment responses to permanent shocks are much bigger than the employment responses to transitory shocks. “Snapshot” health indicators that measure health at a particular point in time may not be able to distinguish transitory health conditions that should not affect employment from permanent health conditions that should (Blundell et al., 2016b).

3.2 Health and income

Table 3 uses data from the NHIS to show how several health indicators vary by age, education and income in the US. Health and income have a strong positive relationship. To get a better sense of this relationship, O’Donnell et al. (2015) decompose the Gini index for income in the US, the Netherlands and China. In each country they regress individual income on age, gender, education, race, region, and self-assessed health, allowing them to express each person’s income as the sum of a predicted value and a regression residual. This allows them to find the marginal impact of each variable on the Gini index, by fixing the variable across all individuals, recomputing each person’s income, and recomputing the Gini index. Using this method for the US, they show that self-assessed health explains (in a statistical sense) 6.5% of income inequality. This is similar to the shares for race and age, but about one-third of that of education. If employment is added to the regression, the share falls to 4.0%, showing that much, although far from all, of health’s relationship with income operates through employment. While these shares are small, most of the income variation is not explained by any covariate whatsoever: not controlling for employment,
self-assessed health accounts for 15% of the explained income inequality in the US. It accounts for similar shares in the Netherlands (15%) and China (10%). O'Donnell et al. (2015) note that their calculations utilize only one measure of health, and a discrete-valued and potentially noisy measure at that. Expanding and improving the set of measures should increase the share of inequality explained by health.

<table>
<thead>
<tr>
<th>Self-Assessed Health</th>
<th>Usual activities limited by chronic conditions</th>
<th>Limited in work</th>
<th>Unable to work</th>
</tr>
</thead>
<tbody>
<tr>
<td>fair</td>
<td>poor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>7.0 (0.10)</td>
<td>12.2 (0.15)</td>
<td>2.8 (0.08)</td>
</tr>
<tr>
<td></td>
<td>2.1 (0.06)</td>
<td></td>
<td>6.2 (0.13)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under 12 years</td>
<td>1.2 (0.11)</td>
<td>7.5 (0.27)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>0.2 (0.03)</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>12-17 years</td>
<td>1.8 (0.17)</td>
<td>9.1 (0.40)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>0.3 (0.09)</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>18-44 years</td>
<td>5.0 (0.15)</td>
<td>5.7 (0.15)</td>
<td>1.7 (0.08)</td>
</tr>
<tr>
<td></td>
<td>1.1 (0.06)</td>
<td></td>
<td>3.2 (0.12)</td>
</tr>
<tr>
<td>45-64 years</td>
<td>11.6 (0.25)</td>
<td>16.3 (0.32)</td>
<td>4.0 (0.15)</td>
</tr>
<tr>
<td></td>
<td>3.9 (0.14)</td>
<td></td>
<td>10.6 (0.27)</td>
</tr>
<tr>
<td>65-74 years</td>
<td>14.5 (0.49)</td>
<td>25.1 (0.62)</td>
<td>7.9 (0.48)</td>
</tr>
<tr>
<td></td>
<td>4.9 (0.30)</td>
<td></td>
<td>11.6 (0.55)</td>
</tr>
<tr>
<td>75 years and over</td>
<td>18.0 (0.60)</td>
<td>42.9 (0.92)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>6.9 (0.41)</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Family income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than $35,000</td>
<td>12.8 (0.25)</td>
<td>21.7 (0.33)</td>
<td>5.1 (0.19)</td>
</tr>
<tr>
<td></td>
<td>4.4 (0.15)</td>
<td></td>
<td>15.4 (0.33)</td>
</tr>
<tr>
<td>$35,000 – $49,999</td>
<td>7.6 (0.30)</td>
<td>12.2 (0.40)</td>
<td>3.1 (0.23)</td>
</tr>
<tr>
<td></td>
<td>2.1 (0.17)</td>
<td></td>
<td>6.1 (0.33)</td>
</tr>
<tr>
<td>$50,000 – $74,999</td>
<td>5.5 (0.26)</td>
<td>9.8 (0.33)</td>
<td>2.7 (0.19)</td>
</tr>
<tr>
<td></td>
<td>1.4 (0.13)</td>
<td></td>
<td>3.8 (0.23)</td>
</tr>
<tr>
<td>$75,000 – $99,999</td>
<td>4.0 (0.27)</td>
<td>8.1 (0.34)</td>
<td>2.2 (0.20)</td>
</tr>
<tr>
<td></td>
<td>1.0 (0.14)</td>
<td></td>
<td>2.6 (0.20)</td>
</tr>
<tr>
<td>$100,000 or more</td>
<td>3.2 (0.17)</td>
<td>6.8 (0.26)</td>
<td>1.4 (0.14)</td>
</tr>
<tr>
<td></td>
<td>0.6 (0.08)</td>
<td></td>
<td>1.6 (0.12)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than a high school diploma</td>
<td>19.8 (0.46)</td>
<td>24.3 (0.55)</td>
<td>4.4 (0.27)</td>
</tr>
<tr>
<td></td>
<td>6.8 (0.31)</td>
<td></td>
<td>14.2 (0.48)</td>
</tr>
<tr>
<td>High school diploma or GED</td>
<td>11.9 (0.30)</td>
<td>17.6 (0.36)</td>
<td>3.7 (0.18)</td>
</tr>
<tr>
<td></td>
<td>3.8 (0.17)</td>
<td></td>
<td>9.4 (0.30)</td>
</tr>
<tr>
<td>Some college</td>
<td>9.2 (0.26)</td>
<td>15.2 (0.33)</td>
<td>3.5 (0.17)</td>
</tr>
<tr>
<td></td>
<td>2.6 (0.14)</td>
<td></td>
<td>7.1 (0.24)</td>
</tr>
<tr>
<td>Bachelor’s degree or higher</td>
<td>4.2 (0.18)</td>
<td>8.6 (0.25)</td>
<td>1.8 (0.11)</td>
</tr>
<tr>
<td></td>
<td>1.2 (0.10)</td>
<td></td>
<td>2.1 (0.13)</td>
</tr>
</tbody>
</table>


Notes: 1Ages 18-69. 2Ages 65-69. 3Ages 25 and older.
3.3 Health and employment

A large body of work shows that health and employment have a positive cross-sectional relationship: see Currie and Madrian (1999) for a review of the older evidence and O’Donnell et al. (2015) for a review of recent studies.

Table 3 shows that as people age, they become more likely to be in fair or poor health, have chronic conditions that limit their activities, or find themselves limited in or unable to work. A rough sense of how this affects labor supply can be found in Figure 3, which shows age-employment profiles for male heads of households. Figure 3 is taken from French (2005), who uses a fixed-effects estimator, corrects for measurement error in health, and uses data from the Panel Study of Income Dynamics (PSID) for the years 1968-1997. Comparing the profiles for (self-assessed) healthy and unhealthy men yields three conclusions. First, individuals in bad health have significantly lower employment rates at most ages. Second, the difference in employment rates between healthy and unhealthy people is especially large for those in their 50s. Third, most of the decline in employment between ages 55 and 70 is explained by declining employment by age, conditional on health, rather than declines in health; health is not the principal driver of retirement. For example, French (2005) finds that the increase in the number of unhealthy individuals from age 55 to age 70, from 20% to 37%, is far smaller than the drop in employment, from 87% to 13%. Even if men in bad health did not work at all, declining health could explain only a fraction of retirement. With the employment profiles shown in Figure 3, bad health can explain only 7 percentage points of the 74 percentage point drop.

Using data from 12 countries, Coile et al. (forthcoming) reach similar conclusions. Applying the approach developed by Cutler et al. (2013), in each country they combine a collection of health indicators into a health index (as in Poterba et al. (2011)) and then compare the employment rates of older people to those of younger people in similar health. Comparisons in all 12 countries suggest that health on average declines with age.
They show that among men aged 50-54, those residing 10 percentile points higher in their country's health (index) distribution are between 2.7% and 7.7% more likely to be working (Table 4). The range for women, 0.6%-6.0%, is similar. However, the declines in health are much too small to explain the declines in employment observed over the same ages.

Several recent studies exploit panel data to better capture health dynamics and control for fixed effects. Bound et al. (1999) conclude that declines in health, as opposed to poor health itself, are important predictors of retirement. In contrast, Disney et al. (2006) find that current and lagged health are both positively related to employment. They attribute the differences between their results and those in Bound et al. (1999) to differences in the time series properties of their respective health measures. Blundell et al. (2016b) show that the process for health contains distinct persistent and transitory components. Employment responds more to permanent than transitory health shocks. Furthermore, the lags of these shocks appear to be important. As a result, the long-run effect of a permanent shock is bigger than the short-run effect.
We should reiterate that these estimates are very sensitive to the measures and methods used. For example, Bound et al. (1999) use a dynamic framework that includes lags of health, and instrument for subjective health measures with objective conditions. They find that bad health reduces employment rates by 55 percentage points for men and 46 percentage points for women. In contrast, McClellan (1998) estimates how the onset of new health conditions affects employment using only objective conditions and contemporaneous changes in health and employment. He finds that major conditions (cancer, heart disease, lung disease) reduce employment 18-26 percentage points, whereas minor conditions (such as hypertension and diabetes) reduce employment only 2-6 percentage points.

3.4 Health and hours of work among workers

Using the same PSID data as Figure 3, Figure 4 shows annual hours of work for employed men. Comparing Figures 3 and 4 shows that bad health reduces both employment and hours worked when employed, but that most of the variation is coming from the employment margin. This should not be too surprising, as most people work full time or not at all. For example, (Blundell et al., 2016a) show among US men aged 60-64, 44.7% do not work at all, 47.4% work at least 1500 hours per year (50 weeks at 30 hours per week), and only 8% work between 1 and 1500 hours per year. More generally, most of the reduction in labor supply associated with retirement takes place along the extensive (employment) margin rather than along the intensive (hours) margin. The prevailing explanation for this finding is that there are fixed costs to work, for both employers and employees, that discourage small amounts of hours (Rogerson and Wallenius 2013, Blundell et al. 2016a).\footnote{Fan (2015) develops a model where “sharp” retirement is driven by habit persistence.}

These fixed costs will be an important element in the model we analyze below.
3.5 Health and wages

Individuals in poor health may be less productive, leading their wages to be lower. Surveying the recent literature, O’Donnell et al. (2015) conclude that bad health leads to modestly lower wages.

Estimating the impact of health on wages is more difficult than estimating its impact on hours or employment, because we can only measure wages for those who work. Because people in poor health are less likely to work than those in good health, it may be the case that the only people in poor health who work are those who receive unusually good wage draws. This would lead us to overstate the wages of the unhealthy and understate the wage gap between the healthy and the unhealthy.
Figure 5 presents estimates of the impact of age and health on wages from French (2005). These estimates are from a fixed-effects regression of wages on a full set of age dummies, and age dummies interacted with health, with corrections for measurement error in health. The figure shows that the wages of those who self report being in bad health are approximately 10% lower.

The fixed-effects estimator identifies the effect of health on wages from the wage changes of those who were in good health one year and bad health the next. This does not fully overcome the selection problem, however, because the potential wages of nonworkers remain unobserved. While the fixed-effects estimator demeans the wages of each individual in the sample, the individual-specific mean is based only on observed wages. The fixed-effects estimator thus measures the wage growth rates of individuals who continue
to work. If wage growth rates for workers and nonworkers are the same, composition bias problems – the question of whether high wage or low wage individuals drop out of the labor market – will not be a problem. However, if individuals with worsening health are more likely leave the market because of a sudden wage drop, then wage growth rates for workers with worsening health will have a greater upward bias than wage growth rates for nonworkers with stable health. This will lead the fixed-effects estimator to understate the wage losses associated with worsening health.

French (2005) addresses this problem by using an estimated structural model of labor supply. This model creates selection since individuals in the model work only when it is in their interest to do so. He finds the wage profiles that, when fed into the model and used to generate artificial wage histories, yield the same fixed-effect regression coefficients for workers as the data. Using this approach, he finds that failing to account for selection leads to a 2 percentage point underestimate of the wage gap between healthy and unhealthy people.

Capatina (2015) adopts a similar strategy, calibrating the wage process that allows the model-simulated wage data for workers to match the cross-sectional wage data from the Current Population Survey. She finds that bad health leads to significant reductions in wages. Her estimates imply that for people under the age of 65 without a college education, wages for those in average and poor health are 19% and 36% lower, respectively, than wages for those in good health. For those with a college education, the corresponding reductions are 14% and 28%.

4 A structural model of saving and labor supply

The reduced form evidence presented in the previous section suggests that bad health is associated with lower labor supply. However, these reduced form estimates often fail to identify the mechanisms by which health affects retirement. Structural models can,
in principle, allow us to quantify the channels – preferences, productivity, and financial incentives – through which health affects behavior. A number of recent studies have made good progress on this front.

To fix ideas, in this section we exposit a fairly standard structural model, namely the model formulated and estimated in French and Jones (2011), to illustrate some of the key channels by which health impacts retirement. This model has been found to fit the data well but is parsimonious enough to make clear many key retirement incentives.

4.1 The model

In this model people choose consumption, labor supply, and whether or not to apply for Social Security. They are allowed to save but they cannot borrow against future labor, private pension, or Social Security income. When making these decisions, they are faced with several forms of uncertainty: survival uncertainty, health uncertainty, wage uncertainty, and medical expense uncertainty.

Consider a household head seeking to maximize his expected discounted lifetime utility at age \( t, t = 30, 31, \ldots, 94 \). Each period that he lives, the individual derives utility from consumption, \( C_t \), and hours of leisure, \( L_t \), according to

\[
U(C_t, L_t) = \frac{1}{1-\nu} \left( C_t^{\gamma} L_t^{1-\gamma} \right)^{1-\nu}. \tag{1}
\]

The quantity of leisure depends on the individual’s labor supply and health:

\[
L_t = L - N_t - \phi_P P_t - \phi_{RE} P_t (1 - P_{t-1}) - \phi_H H_t, \tag{2}
\]

where \( L \) is the individual’s total annual time endowment. Participation in the labor force is denoted by \( P_t \), a 0-1 indicator equal to 1 when hours worked, \( N_t \), are positive. The fixed cost of work, \( \phi_P \), is treated as a loss of leisure. Including fixed costs helps us capture
the empirical regularity that annual hours of work are clustered around 2,000 hours and 0 hours per year. The fixed cost of re-entry, $\phi_{RE}$, limits the extent to which older workers “churn” in and out of the labor market. The quantity of leisure also depends on an individual’s health status through the 0-1 indicator $H_t$, which equals 1 when his health is bad – in other words, there is a fixed cost of being in poor health. Including a time cost of bad health helps us capture the empirical regularity that sick people work less.

Individuals weight future utility using the subjective discount factor $\beta$. Their age of death is uncertain. A worker alive at age $t$ with health status $H$ survives to age $t+1$ with probability $s_{t,H}$. Workers who die value bequests of assets, $A_t$, according to the function $b(A_t)$.

Worker health is also uncertain, following an age-dependent Markov transition matrix. As a matter of notation, let $\pi_{t,H,H'}$ denote the probability (conditional on surviving) that a worker with age-$t$ health status $H$ has age-$t+1$ health status $H'$.

The logarithm of wages at time $t$, $\ln W_t$, is a function of hours worked, age, health status, and the autoregressive component $\omega_t$:

$$\ln W_t = \alpha \ln N_t + W(H_t, t) + \omega_t,$$

where the function $W(N_t, t)$ is the one that fits the wage profile in Figure 5, after controlling for hours worked. The adjustment for hours of work, $\alpha \ln N_t$, reflects the finding in Aaronson and French (2004) and Casanova (2013) that switching from full to part-time work leads to lower wages. The idiosyncratic component of wages, $\omega_t$, follows an AR(1) process. By assumption at time $t$ the individual knows $\omega_t$ but is uncertain about its future values.

The asset accumulation equation is:

$$A_{t+1} = A_t + Y_t + F_t - C_t - M_t, \quad A_{t+1} \geq 0,$$  

22
where $Y_t$ is total post-tax income from all sources. In more detail

$$Y_t = Y \left( rA_t + ys_t + pb_t, W_tN_t, B_t \times ss_t \right),$$

(5)

where: $Y(\cdot)$ is a function converting pre-tax income to post-tax income; $ys_t$ is spousal income, assumed to be exogenous; $pb_t$ denotes benefits from private pensions; $ss_t$ denotes Social Security benefits; and $B_t$ is the 0-1 indicator that equals 1 when the individual has applied for Social Security benefits. The details of how we model private pensions and Social Security can be found in French (2005) and French and Jones (2011).

The variable $M_t$ denotes the individual’s out-of-pocket medical expenditures, the sum of deductibles, co-payments and insurance premia. Medical expenditures depend on the individual’s health insurance, $HI_t$, his health, $H_t$, and residual uncertainty captured by the term $\zeta_t$. The individual’s work decisions may affect his health insurance, but in all other respects both his health and his medical expenditures are exogenous.

Finally, the variable $F_t$ captures government transfers not related to Social Security or universal health insurance. $F_t$ includes means-tested social insurance, such as the Supplemental Security Income and Medicaid programs in the US. Following Hubbard et al. (1994, 1995), such programs are usually modelled as a consumption floor, where government transfers guarantee a minimum level of consumption. $F_t$ can also include payments made under public disability insurance (DI) programs. Since individuals cannot legally work and draw benefits at the same time, being eligible for DI benefits provides a strong work disincentive. However, to be eligible for DI benefits an individual must have a certified work disability, and the approval process is often slow and uncertain. We will use the variable $DI \in \{\text{not applying, under review, receiving benefits}\}$ to denote the individual’s DI status.

Letting $AIME_t$ index Social Security wealth, the individual’s state vector is $X_t =$
In recursive form, the consumer’s problem can be written as

\[ V_t(A_t, \omega_t, H_t, HI_t, \zeta_t, AIME_t, B_t, DI_t) = \max_{C_t, H_t, B_t \geq B_{t-1}, DI_t} \left\{ \frac{1}{1-\nu} \left( C_t^\gamma L_{t-1}^{1-\gamma} \right)^{1-\nu} + \beta s_t H_t \left( V_{t+1}(A_{t+1}, \omega_{t+1}, H_{t+1}, HI_{t+1}, \zeta_{t+1}, AIME_{t+1}, B_t, DI_t) + \beta (1-s_{t+1})b(A_{t+1}) \right) \right\} \]

subject to the equations described above and the laws of motion for \( HI_t, \zeta_t, AIME_t, \) and \( DI_t \). The decision rules are solved recursively, starting with the bequest function at age \( T = 95 \) and working backward. Since there is no closed form solution to the problem, the state variables are discretized into a finite number of points on a grid and the value function is evaluated at those points. With the decision rules in hand, we can then simulate artificial life histories and compare the artificial data to actual data. This allows us to estimate the model through the Method of Simulated Moments (as done by Gourinchas and Parker 2002; Cagetti 2003; and French 2005). In French and Jones (2011), we show that the estimate fits quite well HRS data on employment, hours of work, and assets.

### 4.2 How health affects retirement

An important implication of there being fixed costs to work is that when the net benefit from working is small, small changes in wages or other incentives can lead potential workers to switch between large amounts of work and no work at all, resulting in a high elasticity of labor supply. Conversely, when the net benefit of working is large – positive or negative – the supply of labor is inelastic.

---

8Pension wealth and spousal income are treated as functions of the other state variables and are thus not state variables themselves. Our approximation of pension wealth also requires us to make adjustments to current income, which can be considered a part of \( pb_t \). See French (2005) and French and Jones (2011).
As discussed in French (2005), French and Jones (2012), Karabarbounis (2016) and Blundell et al. (2016a), workers (at least males) are more likely to be indifferent toward work near the ends of their careers. For example, French and Jones (2012, Table 2) show that for the model in French (2005), the labor supply elasticity for a temporary (one-year) wage change rises from 0.36 at age 40 to 1.28 at age 60. Worsening health is only one reason why. As workers age, they may accumulate larger stocks of liquid financial assets or gain greater access to previously illiquid pensions. The tax and pension incentives facing them may be increasingly unfavorable toward work. Their wages may be falling. Most aging workers thus pass from receiving a large net benefit from work to receiving a relatively small benefit from work to facing a large net loss. During the intermediate period, when retirement occurs, their labor supply can be quite elastic. This is one reason why changes in public pensions appear to have such large effects (French and Jones 2012; and Blundell et al. 2016a). This also implies that health conditions that had no effect on employment at younger ages may now induce retirement. Such a dynamic is consistent with Figure 3, which shows that until healthy workers begin retiring in bulk at age 60 the employment gap between healthy and unhealthy workers is increasing with age.

Bad health can discourage work through a number of distinct channels, all of which can be expressed using the model just presented:

1. Preferences. Bad health can raise the marginal utility of leisure relative to that consumption (Capatina, 2015). This is embodied in equation (2), where bad health reduces the agent’s time endowment.

2. Productivity. Bad health can lower workers’ productivity and resulting wage offers. This can be seen in equation (3), where the function $W(H_t, t)$ captures the negative impact of health on wages.

3. Life expectancy. With shorter expected lifespans, individuals in bad health may not need to work as long accumulating financial and pension wealth for retirement. This
effect operates through the term \( s_{t,H_t} \) in equation (6). It also operates through the Social Security system. Individuals who claim Social Security benefits at younger ages receive lower annual benefits. This implies that, all else equal, unhealthy workers are more likely to claim Social Security at younger ages. Because the receipt of Social Security benefits discourages work, through the earnings test (Engelhardt and Kumar, 2014) or income taxation (Jones and Li, 2016) or the relaxation of liquidity constraints (Kahn, 1988), unhealthy workers are also most likely to retire at younger ages. An additional dynamic not captured in the model, where wages are exogenous, is that bad health may discourage the accumulation of human capital.

4. Disability insurance benefits. People in sufficiently bad health may qualify for benefits from firm- or government-based disability programs. This is embodied by the term \( F_t \) in equation (4), which in turn depends on the the disability indicator \( DI_t \). Those receiving benefits have incentives to reduce labor supply through several channels. First, the benefits provide income, allowing individuals to purchase more leisure. Second, in many countries those earning above a certain amount (called the Substantial Gainful Activity level in the US) will lose benefits. Third, in the United States, beneficiaries can receive Medicare or Medicaid health insurance, depending on their income level when working. Here too, excessively high income will trigger the loss of benefits.

5. Medical expenses. Prior to the Affordable Care Act (ACA, “Obamacare”) many US workers could receive actuarially fair health insurance only while they continued to work. In such cases, expensive medical conditions could induce a form of “job lock,” where workers delayed retirement to maintain their health insurance coverage. These dynamics are captured in the dependence of out-of-pocket medical expenditures, \( M_t \), on health insurance, \( HI_t \), and the institutions linking health insurance to work. Alternatively, poor individuals may quality for means-tested public health insurance
(e.g., Pashchenko and Porapakkarm 2013), and DI recipients in the US gain access to public health insurance provided through Medicare (e.g., Kitao 2014a; Kim 2012). These mechanisms, operating through the variables $F_t$ and $M_t$, discourage work.

Most papers consider only a subset of these channels. For example, French (2005) and Capatina (2015) consider four of the five channels but do not account for disability benefits. Kitao (2014a) accounts for DI, but uses very stylized models of demographic transitions and health insurance.

4.3 The determination of health

In the model articulated above, health and employment are related because because health alters wages and/or work incentives. As we noted above, however, much attention has also been given to the possibility that the causality runs at least in part from income to health. Reviewing the literature, O’Donnell et al. (2015) conclude that in developed countries, most with universal health care, the causal effect of income on health is “not proven.” The effect, should it exist, is most likely to operate through child development: “The evidence that childhood health is influenced by economic background and determines adult economic outcomes is persuasive but not yet concrete.”

The canonical economic model of endogenous health determination is that of Grossman (1972), where individuals combine their personal time and health care services to invest in their “health capital.” Quantitative versions of the Grossman model have become increasingly common: recent applications include Yogo (2009), Khwaja (2010), Davis (2006), Halliday et al. (forthcoming), Hugonnier et al. (2012), Scholz and Seshadri (2016), Ozkan (2014) and Cole et al. (2014). Nonetheless, the structural model developed above is consistent with much of the retirement literature in that health is assumed to be exogenous. De Nardi et al. (2016) review the literature on how health care affects health status and conclude that most of the leading studies on this topic find small effects. This suggests
that treating health as exogenous is a reasonable simplification, at least for older individuals, who have established health histories and habits and broad access to public health insurance. Models with both retirement and health investment include Blau and Gilleskie (2008), Fonseca et al. (2009), Li (2015) and Pelgrin and St-Amour (2016).

Table 3 also shows a strong negative relationship between health and education. Looking at the US, the UK and France, Coile et al. (forthcoming) document that people in the lower educational attainment quartiles are more likely to have fair or poor self-assessed health than those at the top. Bound et al. (2014) find a similar educational gradient for mortality rates in the US. Disentangling the interactions between health, education and income, if even feasible, lies well outside the scope of this paper: useful reviews include Deaton (2003), Cutler and Lleras-Muney (2012), and O’Donnell et al. (2015). In the context of the structural model developed above, it is important to recognize that education can help forecast both health and wages, and may signal unobserved differences in patience, risk aversion, self-control, or attitudes toward work. Many structural models of retirement include education in their state vector.

5 Recent structural evidence on how health affects retirement

5.1 Preferences, productivity, and survival

Capatina (2015) develops a life-cycle model of US workers where health affects time endowments, labor productivity, mortality and medical expenses. She finds that the main channels by which health affects earnings are through its effect on productivity and

---

9Blau and Gilleskie (2008), Fonseca et al. (2009) and Li (2015) do not discuss how their results would change if medical expenses were treated as exogenous. Pelgrin and St-Amour (2016) briefly discuss the role of endogenous medical spending but conduct no quantitative assessments. Theoretical analyses that embed retirement decisions within the Grossman framework include Wolfe (1985) and Galama et al. (2013).
through time lost to bad health. These mechanisms are much stronger for individuals without a college education. For this group, eliminating the productivity losses due to bad health would lead labor supply to increase by 7.4% (Table 7). Eliminating the time cost of bad health would lead labor supply to increase by 3.3%.

Capatina (2015) finds that removing the lifespan reductions associated with bad health would lead to large increases in saving, especially at older ages, but would have relatively little effect on labor supply. In Capatina’s experiment, the incidence of bad health, and the productivity losses and time costs associated with it, remain at their baseline levels. As a result, increased lifespans are not matched by “proportional” (Bloom et al., 2014) increases in potential lifetime earnings, and their effects on labor supply are modest. Capatina (2015) also holds fixed the parameters of the Social Security system.

Haan and Prowse (2014) construct and estimate a similar model for single German workers and use it to forecast how they will respond to increased longevity. In contrast to Capatina (2015), Haan and Prowse (2014) allow increases in longevity to be accompanied by increases in pension age thresholds and employment opportunities. The last modification is particularly important because their estimates show that job offer rates decrease rapidly and involuntary separations rise rapidly as workers approach the full public pension age. They find that if lifespans increase by the expected 6.4 years, but there are no other changes, years of work after age 40 increase by only six months. With such a limited employment response, increased longevity will render the German pension system insolvent. In contrast, if the full pension age (and employment opportunities) increase by 3.76 years, years of work will increase by almost three years. This will be sufficient to keep the pension system in balance.

An important feature of mortality risk is that not only mortality but mortality rates are uncertain. In many structural models this mechanism is captured by allowing mortality rates to depend on health, which is itself stochastic. This linkage has subtle but important effects. For example, Reichling and Smetters (2015) examine how health-related mortal-
ity risk reduces the demand for annuities. Because a bad health shock simultaneously raises a person’s expected (current) medical expenses and lowers her expected lifespan, it simultaneously lowers an annuity’s actuarial value and raises the need for liquid wealth. Imrohoroglu and Kitao (2012) make a related point in the context of Social Security. Recall that that workers who start receiving Social Security benefits at older ages receive higher annual benefits. While the optimal claiming age depends in large part on a worker’s expected lifespan, uncertain mortality introduces precautionary considerations. Workers may delay claiming to guard against the possibility that they receive a positive longevity shock in the future.

5.2 Disability Insurance

Disability insurance is an important pathway to early retirement, especially because disabling conditions are most likely at older ages. Modeling disability insurance is often challenging, because the benefit approval process can be complicated and uncertain. The most common reasons for applying for benefits in the US are musculoskeletal (such as back) and mental health problems, both of which are difficult to diagnose. Furthermore, in order to be eligible for benefits, the individual must be out of the labor force, including when applying for benefits.

Most structural studies of disability insurance emphasize the modelling and measurement of health. Bound et al. (2010) and Iskhakov (2010) treat health as a latent variable. Treating “true health” as unobserved affects the way in which eligibility for disability benefits is determined within the model, and how health is measured when estimating the model from the data. Bound et al. (2010) find that when health is treated as a latent variable, the estimated behavioral effects of health are smaller than when health is measured in more conventional ways. As we discussed in section 3.1, reduced form estimates of the effects of health are sensitive to how health is measured. Structural studies such
as Bound et al. (2010) and Iskhakov (2010) could help determine which health measures are best. Laun and Wallenius (forthcoming) use a calibrated model with latent health to examine how the effects of disability insurance differ across countries.

Low and Pistaferri (2010) and Benitez-Silva et al. (2011) add savings to the model, so that self-insurance through savings can serve as an alternative to disability benefits. Kitao (2014a) calibrates a model with uncertain medical spending as well as disability. This is an important extension because beneficiaries of the US Social Security Disability Insurance (SSDI) program also become eligible for Medicare health insurance. Kitao (2014a) finds that if the Medicare benefits were removed, SSDI recipiency would fall by 30%. Kim (2012) considers similar issues in an estimated model. He finds that the Medicare component of SSDI reduces the employment rate of men aged 23-62 by 0.7 percentage points.

French and Song (2016) estimate a dynamic programming model of employment and the disability application process, fitting the model to administrative data on flows into the SSDI program, appeal rates of those denied, and the labor supply of those allowed and denied benefits during the application process. They account for the fact that by law individuals cannot work and draw benefits at the same time. Furthermore, those applying for benefits or appealing a benefit denial will receive a “technical denial” if they work. French and Song (2016) account for the fact that many of those who apply for benefits are initially denied but remain out of the labor force so that they can continue to appeal their rejection. They find that failing to account for these dynamic incentives would significantly understate the true work disincentive of disability insurance.

Jacobs (2015) builds on the observation that the risks of disability vary greatly by occupation. She builds a life-cycle model where individuals consider disability risks, and the insurance provided by disability insurance, when choosing occupations. She finds that the loss from eliminating disability insurance is equivalent to a 7-percentage point reduction in lifetime consumption for white collar workers and a 12-percentage point
reduction for blue collar workers.

Some studies consider potential reforms to the US disability insurance system. Under current rules, individuals with earnings above a (low) ceiling lose all their disability benefits. Benitez-Silva et al. (2011) consider a reform that allows disability beneficiaries to keep 50 cents of each dollar they earn. The impact of this reform depends on whether beneficiaries believe that earnings make them more likely to undergo continuing disability reviews. They also conclude that the reform would not significantly increase the number of beneficiaries though “induced entry.” Yin (2015) finds that introducing partial disability insurance would lead to financial savings and higher welfare for the disabled.

Milligan and Wise (2012) calculate disability recipiency rates at age 64 for eight countries and find rates ranging from less than 10% to over 35% (Figure I.1). These differences may be due to cross-country variation in the disability programs themselves or to variation in the programs for which disability insurance can substitute. A number of studies consider interactions between disability insurance and other transfer programs. Considering the Netherlands, Heyma (2004) finds that restrictions on disability insurance would lead more workers to use unemployment insurance, leading employment at age 60 to increase by only 1.7 percentage points. Iskhakov (2010) finds that some Norwegian workers would respond to the elimination of the early retirement option by acquiring partial disability benefits (available in Norway) but few would acquire full benefits. Laun and Wallenius (2015) use a calibrated model to assess the effects of recent reforms to the Swedish pension system. They find that these reforms have little effect on the use of disability insurance, as disability benefits were more generous than standard pensions even before the pension reforms.

Kitao (2014a) finds that disability insurance amplifies the effect of reforms to the US Social Security and Medicare systems. In particular, reforms that increase the need for retirement wealth encourage workers to stay in the labor force rather than take early retirement through disability insurance. Nonetheless, more analysis is needed. For ex-
ample, the data show that increases in the normal retirement age tend to increase entry into disability insurance. It would be useful to have a model to better assess the mechanisms behind these changes and to better assess the optimality of increasing the normal retirement age.

5.3 Medical spending risk and health insurance

In the US a major risk is the possibility of facing high medical spending when uninsured. This risk can impact retirement through multiple channels.

First, it is potentially an important driver of savings (Palumbo 1999, De Nardi et al. 2010, Kopecky and Koreshkova 2014, De Nardi et al. 2016). This in turn potentially affects labor supply, as workers may seek higher earnings to build up their precautionary savings.

Second, it implies that health insurance may have significant effects on labor supply. In the US the majority of households younger than 65 receive health insurance through their employers. Prior to the ACA, this was the only group health insurance available to most of them, and the insurance was available only while they continued to work.\footnote{Employees can buy health insurance from their former employers at (unsubsidized) group rates via the “COBRA” provision for up to 18 months after leave their jobs.} For this reason many individuals may have been working not just for their salaries, but also for the health insurance provided by their employers. This relationship changes abruptly at age 65, when almost everyone qualifies for Medicare health insurance provided by the US government. Once individuals become eligible for Medicare, the health insurance incentive for work largely vanishes.

Consistent with this view, individuals who would lose their health insurance if they retired before age 65 tend to remain at their jobs about six months longer than those who can retain their coverage post-retirement. This has been documented by Madrian (1994), Rust and Phelan (1997), and Blau and Gilleskie (2001), among others. Rust
and Phelan (1997) and French and Jones (2011) show that the differences in employment among health insurance types is largely explained by the greater tendency of individuals whose health insurance is tied to their jobs to remain in the labor market until they are eligible for Medicare at age 65. Figure 6, taken from French and Jones (2011), shows that those whose health insurance is tied to their jobs have high job exit rates at age 65, whereas those with access to post-retirement “retiree” coverage tend to leave at 62, when they can first receive Social Security benefits. Those with no employer-provided health insurance have less pronounced spikes at both ages.

Structural work on this issue began with Gustman and Steinmeier (1994). They estimate that retiree coverage reduces years in the labor force by 0.1 years, a much smaller effect than the reduced form estimates. One likely reason why Gustman and Steinmeier (1994) find small effects is that they assume that medical expenses are certain, and they
find that the average employer contribution to health insurance is small. However, if employer-provided coverage allows workers to reduce their exposure to medical expense risk, and actuarially fair insurance is not available in the non-group market, workers may value their insurance benefits well above the benefits’ actuarial cost. Rust and Phelan (1997), who allow medical expenses to be risky, find labor supply responses to employer-provided insurance that are much larger. Blau and Gilleskie (2006) allow couples to insure through joint labor supply, while Blau and Gilleskie (2008) reduce medical spending risk by allowing medical spending to be endogenous. Perhaps not surprisingly, the effects of health insurance that they report are smaller than those of Rust and Phelan (1997), albeit larger than those of Gustman and Steinmeier (1994).

Rust and Phelan (1997) and Blau and Gilleskie (2006, 2008) all assume that an individual’s consumption equals his income net of out-of-pocket medical expenses and thus ignore his ability to self-insure through saving. French and Jones (2011), who allow workers to accumulate assets, find health insurance effects that are larger than those of Blau and Gilleskie (2006, 2008) but smaller than those of Rust and Phelan (1997). French and Jones (2011) find that Medicare was about as important as Social Security in determining retirement for the cohort that turned 65 in the late 1990s. In their model raising the Medicare eligibility age from 65 to 67 leads individuals to work an additional 0.074 years over ages 60-69, while eliminating two years’ worth of Social Security benefits increases years of work by 0.076 years. All of these studies match better the reduced form relationships between health insurance and retirement than Gustman and Steinmeier (1994), showing the importance of catastrophic medical expense risk for understanding retirement.

Pelgrin and St-Amour (2016) estimate a life-cycle model of labor supply, saving and health capital accumulation with health insurance. They consider the many decision margins that health insurance might impact, including saving, health and labor supply.

---

11van der Klaauw and Wolpin (2008) and Casanova (2011) also consider the importance of spouses in structural models with medical spending but do not focus on the effects of health insurance.
Fonseca et al. (2009) use a similar model to analyze the causes of the growth in life expectancy and medical spending between 1965 and 2005. They conclude that increases in health insurance generosity and improvements in medical technology play roughly even roles in the expansion of medical spending. Technology improvements explain most of the increase in longevity.

A number of recent papers consider how the Obamacare health care reforms might affect retirement. French et al. (2016) estimate a structural model that accounts for two important provisions of the ACA: (1) the expansion of the Medicaid health insurance program for low-income households younger than 65, including the elimination of the wealth test; (2) the reforms to the private nongroup insurance market, including the introduction of income-based subsidies. Both reforms should strengthen the ability of households to maintain health insurance coverage after leaving the jobs, which should affect their retirement decisions. Li (2015) analyzes the ACA reforms in a general equilibrium model with endogenous health. She pays particular attention to the interactions between the ACA reforms and SSDI, finding that expanded access to health insurance reduces the rate at which people develop disabilities, reducing the net cost of the ACA.

Hansen et al. (2014) develop a life-cycle model where individuals can “buy into” Medicare, the principal public health insurance program for elderly Americans. Medicare buy-in has been advocated by Democratic presidential candidates both past (Al Gore) and present (Hillary Clinton). Hansen et al.’s (2014) calibrations suggest that the market for Medicare buy-in will not develop without subsidies, because of adverse selection driven by the option of private insurance. The subsidies needed to make the program effective among the target population (ages 55-64), however, would require only a small increase in labor taxes.

Many US households receive health insurance through the Medicaid program, which is means-tested. In principle, means-tested health insurance, which is more targeted than universal coverage, could be very effective. However, it also distorts incentives to work,
save, and consume medical care. Pashchenko and Porapakkarm (2013) and Pashchenko and Porapakkarm (2015) find that the work disincentives of Medicaid are significant and costly. Mulligan (2013) argues that because the subsidies for private health insurance provided in the ACA decrease with household income, they too discourage work.

Interestingly, preliminary evidence suggests that the Obamacare reforms had at most a modest effect on retirement patterns. For example, Levy et al. (2015) compare states that offered Medicaid expansions with those that did not. They find the expansions appear to have had virtually no effect on retirement. Whether these results suggest that Medicaid has no impact on retirement, or only that we can rule out extremely large work disincentives from the Medicaid expansion, remains to be seen. Evidence from the Oregon Health Insurance Experiment also suggests a very small Medicaid work disincentive. Baicker et al. (2014) find that access to Medicaid via the Experiment lottery reduced employment rates by a statistically insignificant two percentage points.

6 Conclusions and Suggestions for Future Work

Over the past century the developed world has enjoyed an increase in healthy lifespans accompanied by a move toward earlier retirement. These trends, coupled with reduced fertility and increasing medical costs, have led to persistent fiscal imbalances. Many countries have responded with policy reforms intended to promote later retirement. The academic literature on health and retirement contains important lessons for those seeking reform.

First, the long-term decline in retirement has not been driven by health. More plausible explanations are the wealth effects of higher wages, incentives generated by public pension and disability insurance programs, or declines in the cost of goods complementary to leisure.

Second, the cross-sectional data show that unhealthy people retire earlier. It is much
less clear whether in general bad health is a compelling force, a tipping factor in an environment already conducive to early retirement, or a gateway to disability benefits. Comparisons both reduced-form and structural suggest that much of the cross-country variation in retirement comes from differences in public pensions and disability insurance. The increases in labor supply that have followed recent pension reforms suggest the same thing.

The myriad of ways in which bad health can affect retirement – preferences, productivity, mortality expectations, the need for health insurance (in the US) – along with overlapping and interacting government programs, provide a showcase for careful structural analysis. A particularly promising topic is the intersection of old-age policy reform and income- and education-based differences in health. Chetty et al. (2016) show that during the period 2001-14, the life expectancy at age 40 for men in the top 1% of the US income distribution was 14.6 years longer than for those in the bottom 1%; the gap for women was 10.1 years. Waldron (2007), De Nardi et al. (2010) and Pijoan-Mas and Ríos-Rull (2014) report similar results for older individuals. Jacobs (2015) shows that blue-collar workers are much more likely to suffer disabilities. This raises a host of compelling questions. For instance, is the US Social Security system progressive? Although the formula converting lifetime earnings to annual benefits is progressive, the tendency of richer individuals to receive longer benefit streams is regressive. Other policies and proposed policy reforms deserve similar attention.

---

12Recent quantitative analyses include Goda et al. (2011) and National Academies of Sciences, Engineering, and Medicine (2015).
References


Halliday, Timothy J., Hui He, Lei Ning, and Hao Zhang, “Health Investment over the Life-Cycle,” forthcoming.


Li, Yue, “The Affordable Care Act and Disability Insurance,” March 2015. Mimeo (available at SSRN 2448955).


