Using the Health and Retirement Study for Disability Policy Research: A Review

Jody Schimmel Hyde and David C. Stapleton
Received April 21, 2016

Abstract

The Health and Retirement Study (HRS) is a preeminent data source for research related to the experiences of workers nearing retirement, including the large share of those workers who experience a health shock or disability onset after age 50. In this article, we highlight key information collected from HRS respondents that benefits disability policy research and the body of knowledge that has resulted from this information. Our main goal is to identify from this research experience potential improvements in data collection and documentation that would further strengthen the HRS as a data source for disability policy researchers.

I. Introduction

The purpose of this manuscript is to assess the strengths and limitations of the information in the Health and Retirement Study (HRS) for purposes of research that addresses questions of policy significance related to disability and participation in the Social Security disability program and other programs. We begin by providing a brief overview of some of the key aspects of disability onset for older workers as they near retirement to set the stage for the remainder of our discussion. We then provide a brief overview of our assessment of the primary strands of literature related to disability onset among those prior to retirement age.

The primary focus of our analysis is on three areas of data collection and measurement that are desirable for existing and future research using the HRS to be able to answer pressing disability policy questions. These areas include data related to: (1) the measurement of disability and disability onset; (2) employment and staying employed following disability onset; and (3) the receipt of federal disability benefits. As we will document, in each of these areas, the HRS has significant advantages over other nationally representative data sources for studying the population nearing retirement. At the same time, we provide suggestions for ways in which data in these areas might be improved in future survey years.

A. Policy Background

As workers age, it becomes increasingly likely that they will experience a health shock or the onset of a new work-limiting health condition. Eight years after adults in the HRS were first interviewed (when ages 51-61), more than 80 percent had experienced the onset of at least one chronic condition, and for a third of those, the condition was major (cancer, heart condition, stroke, or lung disease) (Smith 2005). For many of these, onset occurred in their 50s. Similarly,
among 51-55 year olds in the same cohort, one-quarter reported having a health condition limiting their ability to work at least once by age 62 (Johnson et al. 2007).

The onset of new health and functional limitations can have important consequences on employment; numerous studies have documented lower rates of employment among those with disabling conditions (among others, Stapleton and Burkhauser 2003; Autor and Duggan 2003). Interruptions in labor market participation are especially important in the years prior to retirement, as workers have reached the peak of their earnings trajectories and are accumulating assets to support them after they retire. Interruptions in labor force participation, especially early withdrawal from the labor market and premature retirement can mean reduced income and wealth for the remainder of their lives. Among workers who experienced the onset of a new work-limiting health condition after age 50 but before retirement, earnings fell dramatically, resulting in higher poverty rates among those with such conditions relative to comparable adults without such conditions (Schimmel and Stapleton 2012). In addition to lost income and reduced wealth from not working, exiting the labor force may often mean sacrificing employer-provided health insurance prior to the availability of public coverage.

Because of these financial consequences of early labor force withdrawal and in response to macroeconomic concerns about the retirement of the Baby Boom generation, there has been an increased policy focus in recent decades on keeping older workers in the labor force. Remaining in the labor force often requires new skill acquisition or accommodations from one’s employer including physical modifications, changes in job requirements, or more flexibility in working hours. Some workers may need modest financial support to remain working, although the availability of such support is very limited. Because the bulk of public support goes to those who stop working, instead, the support system has the effect of creating work disincentives (Schimmel Hyde and Stapleton 2010).

For those who are unable to remain working due to chronic and significant impairments prior to reaching full retirement age (FRA, currently age 66), the Social Security Administration’s (SSA) disability programs may offer income support. Workers who experience a significant physical or mental impairment expected to last for at least twelve months or result in death are eligible for benefits from the Social Security Disability Insurance (SSDI) program. Eligibility for SSDI is based on meeting the program’s work history requirements and being determined to be unable to engage in substantial gainful activity (SGA). The SSDI program is a key support for many with disabilities approaching retirement; nearly two in three of the 869,000 awards to SSDI disabled workers (those with sufficient work history to qualify for benefits on their own) in 2013 were for applicants ages 50 to FRA (SSA 2015). The average SSDI benefit paid to those who receive an award is modest; $1,114 (per month) for those ages 50-54 in 2013,

---

1 SSDI beneficiaries with meager income and assets may also be eligible for Supplemental Security Income (SSI), also administered by SSA.
rising to $1,318 for those ages 65-FRA (SSA 2015). In addition, new SSDI beneficiaries must wait 24 months after eligibility for SSDI to qualify for Medicare benefits. Based on data from before implementation of the Affordable Care Act (ACA) Medicaid expansions and Health Insurance Exchanges in 2014, many new awardees were without health insurance (Livermore et al. 2010).

**B. Key Areas of Disability Policy Research in the HRS**

Tracking the causes and consequences of disability onset in the pre-retirement years requires data collection that spans many years. Disability evolves over time, occurring as a result of a single health shock or the worsening of several chronic and co-occurring health conditions. Assessing the impact of disability on economic outcomes and benefit claiming behavior requires knowledge of the status of the household prior to the onset of a work-limiting health condition and for many years thereafter. The HRS, with its long panel structure and depth of information on health status, functioning, income, assets, and benefit receipt has become the go-to source for research related to the onset of disability, the claiming of Social Security disability benefits, and the interplay between federal disability and retirement benefits. Data has been collected every two years over twelve waves spanning nearly a quarter of a century, long enough to observe the original panel of 51-61 year olds as they ended their careers, retired, and now for many, are facing complex end of life decisions. The replenishing of the panel every six year to be representative of new cohorts of 51-56 year olds offers an opportunity to compare the experiences of four distinct cohorts as they approach retirement age. For two of these cohorts, enough data has been collected to follow respondents from the peak of their careers to the claiming of Social Security and Medicare.

The body of research related to disability using the HRS is quite broad, with many studies focused on functional limitations among those of retirement age. We focus our analysis on individuals prior to retirement age—the youngest members of cohorts in the HRS (age 50 to FRA). Among this age group, the most significant body of work that would not have been possible without the HRS is in two research areas:

- **Understanding the impact of disability onset on the financial well-being of older households.** Studies in this area have sought to understand how households adjust after the onset of new health events (for example, Coile 2004; Smith 2005) and disability onset (for example, Dushi and Rupp 2013; Stapleton et al. 2008; Schimmel and Stapleton 2010 and 2012; Johnson et al. 2006, 2007 and 2010). These studies relied on the long panel nature of the HRS to identify these events, document the personal situations and household finances after such events, and analyze individual and household responses.

- **Understanding the relationship between disability (or rather, health and functioning) and retirement, including the interplay between Social Security disability and retirement benefits.** The HRS has been used extensively to understand
household behavior, especially with regard to the incentives of public programs such as SSR. Multiple studies have considered the interplay between disability and the claiming of retirement benefits, especially in the years prior to FRA. These studies have shaped the knowledge base surrounding how health and functioning affect the decision to and timing of retirement, and have documented that early retirees often look much more like those with disabilities than those who retire at older ages (for example, Belbase et al. 2015; Bound et al. 1999; Bound and Waidmann 2010; Dwyer and Mitchell 1999; Li and Maestas 2008; McGarry 2004; Munnell et al. 2015).

II. Measuring Disability and Disability Onset in the HRS

Disability is dynamic and multifaceted, reflecting both a person’s physical and mental health as well as the environment in which they live. Two main conceptual models of disability have guided the collection of data in national data sources. Nagi’s (1965) study and later work by Verbrugge and Jette (1994) moved away from the medical model of disability, recognizing that functional limitations are the consequence of a constellation of medical, personal and environmental characteristics. More recently, the World Health Organization’s (WHO) International Classification of Functioning, Disability and Health (ICF) model of disability, which puts these ideas in a formal structure, has been adopted by many researchers, administrators and practitioners. The ICF offers a classification scheme and comprehensive nomenclature for disability and its antecedents, covering biological, individual, and environmental domains (WHO 2002).

Translating the ICF definition of disability to survey data collection is challenging because of its complexity (Hendershot 2002) and because the definition of disability varies widely across the dozens of public programs that exist to support individuals with disabilities (Mashaw and Reno 1996). Beginning in 2008, the United States Census Bureau adopted a six-question sequence (6QS) to assess disability status in the American Community Survey (ACS) (Weathers 2005; Brault et al. 2007). The 6QS has been adopted as the disability sequence in the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC) in 2009 and was included in the Survey of Income and Program Participation (SIPP) as a topical module in 2010 (Burkhauser et al. 2014; Brault 2013; McMenamin and Hipple 2014; Brandt et al. 2014). In addition, the Department of Health and Human Services (DHHS) adopted the 6QS as the data collection standard for measuring disparities between those with and without disabilities in response to a mandate in Section 4302 of the Affordable Care Act (DHHS 2011).

The HRS has not formally adopted the 6QS for assessing disability. In Table 1, we highlight the closest questions we were able to find in the 2014 core HRS. In most cases, these questions deviate fairly significantly from the comparable question in the 6QS. Further, the questions are spread throughout the survey rather than in sequence. The HRS could potentially adopt the 6QS, in place of some or all of these questions, but doing that alone would create a very undesirable
seam for comparisons across waves or panels. Another option would be to maintain the existing questions and add the 6QS as a topical module in one or more future waves of the HRS, to create a benchmark that can be compared to existing measures. That could lead to a revamping of the HRS disability questions thereafter, and produce the information necessary to span any seam that the change would create. Alternatively, the HRS could continue to use the existing questions, with occasional use of the 6QS to support benchmarking.

Table 1. 6QS Used in Many Nationally Representative Surveys to Assess Disability and Closest Comparable Question in the 2014 Core HRS

<table>
<thead>
<tr>
<th>6QS</th>
<th>Closest identified measure in the 2014 core HRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you deaf or do you have serious difficulty hearing?</td>
<td>Do you ever wear a hearing aid? Is your hearing excellent, very good, good, fair, or poor using a hearing aid as usual?</td>
</tr>
<tr>
<td>Are you blind or do you have serious difficulty seeing even when wearing glasses?</td>
<td>Is your eyesight excellent, very good, good, fair, or poor using glasses or corrective lenses as usual?</td>
</tr>
<tr>
<td>Because of a physical, mental, or emotional condition, do you have serious difficulty concentrating, remembering, or making decisions?</td>
<td>During that same two-week period, did you have a lot more trouble concentrating than usual?</td>
</tr>
<tr>
<td>Do you have serious difficulty walking or climbing stairs?</td>
<td>Because of a health problem do you:</td>
</tr>
<tr>
<td></td>
<td>…have any difficulty with walking several blocks?</td>
</tr>
<tr>
<td></td>
<td>…with walking one block?</td>
</tr>
<tr>
<td></td>
<td>…with climbing several flights of stairs without resting?</td>
</tr>
<tr>
<td></td>
<td>…with climbing one flight of stairs without resting?</td>
</tr>
<tr>
<td>Do you have difficulty dressing or bathing?</td>
<td>Because of a health or memory problem do you:</td>
</tr>
<tr>
<td></td>
<td>…have any difficulty with dressing, including putting on shoes and socks?</td>
</tr>
<tr>
<td></td>
<td>…have any difficulty with) bathing or showering?</td>
</tr>
<tr>
<td>Because of a physical, mental, or emotional condition, do you have difficulty doing errands along such as visiting a doctor’s office or shopping?</td>
<td>(Because of a health or memory problem do you have) any difficulty with shopping for groceries?</td>
</tr>
</tbody>
</table>

The breadth of information available in the HRS, despite the absence of the 6QS, has allowed researchers the flexibility to define disability in a number of ways, often targeted to the age of the population of interest. Researchers interested in studying the population who might be at risk for exiting the labor force following disability onset have often used a single measure of whether one’s health limits the type or amount of work a person can perform (Choi 2003; Johnson et al. 2007; Heiss et al. 2008; Kapteyn et al. 2008; Schimmel and Stapleton 2012; Schimmel Hyde and Stapleton 2010).² Others have used this question in conjunction with other studies such as Charles (2003) and Meyer and Mok (2013) have used similar measures for younger populations.

² Other studies such as Charles (2003) and Meyer and Mok (2013) have used similar measures for younger populations.
more objective measures of health, described below (Altinag et al. 2012; Dushi and Rupp 2013; Johnson et al. 2006; Johnson et al. 2010; Loprest et al. 1995; Verbrugge et al. 1999). This question has been available in every wave of the HRS and is available even for those who are above FRA. The strength of this measure is in its simplicity; a single question to capture disability as defined by the respondent. It has been found to be correlated with objective health measures such as reported chronic conditions and functional limitations using other survey data (Bound 1991; Burkhauser and Daly 1996). In addition, it is similar in spirit to the medical eligibility criteria for SSDI and SSI, which are based in part on one’s inability to engage in substantial work activity. Burkhauser at al. (2012) also have found that this type of question elicits positive responses from many SSDI and SSI beneficiaries who do not respond positively to the 6QS; the 6QS alone elicits positive responses from about two thirds of beneficiaries, versus almost 90 percent with the addition of the a work limitation question.

Yet, work-limiting health conditions as a proxy for disability has been questioned for a variety of reasons (Burkhauser et al. 2002, Benitez-Silva et al. 2004). For example, it has been posited that the measure is biased and endogenous; those who have decided to not work may justify their decision by responding that they have a health condition that limits their work activity (Kreider 1999). Other studies have not found evidence supporting the claim of justification bias of self-reported work limitations and other health measures, at least with respect to retirement (Dwyer and Mitchell 1999; McGarry 2004). It also has been suggested that the work-limitation measure may not capture well the population that likely will claim disability benefits nor align with other measures of significant health or functional limitations. Of particular note is that there is a fair amount of variability in the report of work limitations across waves, meaning that health conditions referenced may not lead to significant nor permanent limitations, as SSDI eligibility requires (Schimmel and Stapleton 2012; Kapteyn et al. 2008; Choi 2003). Possibly more relevant for HRS respondents nearing retirement than those younger groups, answers to the work limitation question alone are not definitive about why the person stopped working, nor can they identify the myriad reasons why the person has been unable to return to work (skills deterioration, transportation issues, lack of employers willing to offer a position, etc.).

Numerous researchers, particularly those studying disability of the oldest old, have used the large battery of questions about functional status to define disability. Some studies have used the presence of any limitations in activities of daily living (ADLs) and instrumental activities of daily living (IADLs), while others have paid attention to the number of ADL and IADL difficulties or certain ones (Bowen and Gonzalez 2010; Chan et al. 2011; Choi et al. 2015; Cigolle et al. 2007; Dunlop et al. 2007; Freedman et al. 2004a; Freedman et al. 2004b; Freedman et al. 2008; Latham 2012; Smith et al. 2005; Sturm et al. 2004; Wolf et al. 2015). Limitations in these activities tend to occur relatively infrequently in pre-retirement age groups, and those with multiple limitations would be among the most likely to report having a health condition limiting one’s ability to work, meaning that these functional questions would be extraneous.
Several studies have used data collected in the HRS on health conditions to assess the impact of health shocks on employment and income in the years leading up to retirement (Bound et al. 1995; Bound et al. 1999; Coile 2004; Coile and Milligan 2007; Dushi and Rupp 2013; Johnson et al. 2006; Johnson et al. 2007; Smith 2005). In general, those studies have sought to understand individual and household responses to changes in health prior to retirement, and have not focused on work disability per se. For example, Smith (2005) assessed the onset of new major and minor health conditions, where major was defined to include cancer, heart condition, stroke, and lung disease. While certainly the onset of those conditions can be life-changing, one significant limitation of using these measures as a proxy for disability is that depending on severity, they may not pose any significant limitation in daily activities or work, especially if responsive to treatment (McClellan 2008). The HRS does not query respondents about severity. Inconsistent responses across waves to questions phrased as “has a doctor ever told you that you have” a given condition are often assumed to be respondent error, but may also signal transitory health conditions that are not very severe. As a result, studies that assess the impact of health shocks, while highly informative, may be considered distinct from similar studies assessing disability onset using health-related work limitation. Though a relatively narrow topic compared to broader measures of disability onset, two studies used the HRS to look at the predictors and consequences of occupational injuries (Zwerling et al. 1996; Zwerling et al. 1998).

One particular innovation in the HRS and its sister studies abroad is the development of vignettes in the HRS that allow the benchmarking of disability in the US relative to other countries. These vignettes describe hypothetical individuals’ health conditions or functional limitations, then ask respondents to assess the extent to which the hypothetical individual is work disabled. Because these questions have been included in the HRS as well as studies in other countries, researchers have been able to use them to understand the role that social norms play in determining reports of work disability. These vignettes have been documented to be a novel way to merge subjective and objective information about disability and functioning (Van Soest et al. 2007). Kapteyn et al. (2007) used these vignettes to document that respondents in the Netherlands were much more likely to consider given hypothetical workers to be disabled than respondents in the HRS, thus helping to explain why self-reported work disability is much higher in the Netherlands than in the U.S. Other work by Kapteyn et al. (2008) and Banks et al. (2009) used vignettes to explore the relationship between pain—which can be highly subjective—and reports of work limitations in static and dynamic contexts.

An important point that pertains to any of the selected measures of disability is that the length of the HRS panel makes it an exceptional source of information about the evolution of health, functioning, and disability. No other nationally representative data source in the United States contains panel data with sufficient sample size to explore trajectories in health and disability in the pre- (and post-) retirement years. Being able to study changes in disability over...
time—with consistently collected questions—is a key advantage to using the HRS to study trajectories near and after retirement. One important implication of this for HRS staff, however, is that it adds importance to the avoidance of data collection errors. The consequences of errors in a longitudinal survey are illustrated by an error in the 2004 HRS wave. There was a change in the skip logic for the question about work-limiting health conditions whereby respondents who had reported such a limitation in 2002 were not asked about the presence of such a limitation in 2004. This difference relative to all other waves in several cases has meant that researchers have had to either impute information in that wave, skip it altogether, or end longitudinal analysis in the wave prior.

III. Using the HRS to Study Issues Related to Employment and Staying Employed Following Disability Onset

The HRS collects detailed information about employment, income, and assets that facilitate studying labor force decisions as well as the ways in which households cope with changes in labor force participation resulting from disability onset. We begin by highlighting topic areas that are particularly useful for studying the employment of individuals with disabilities, then turn to highlight other questions related to financial well-being.

The HRS contains a wealth of questions related to employment: work for pay, hours worked, wages, other compensation, firm size, industry, and occupation that can be tailored to many research purposes. In addition, questions related to volunteer work allow for an assessment of social engagement, even if not in paid employment (Zedlewski 2007). There are also several sets of questions in the HRS that are of particular interest for understanding the relationship between disability and employment. The ability of an employee with a disability to remain working may depend critically on the requirements of his or her job, the flexibility of an employer to respond to declining or changing abilities, and the availability of other supports. These questions have been exploited with some frequency for studying the decision to retire, but to our knowledge, few have connected these questions directly to the onset of disability or the decision to seek disability benefits.

- **Job characteristics.** A series of questions in the HRS employment section asks employed respondents to assess whether their current job requires: physical effort; lifting heavy loads; kneeling, stooping, or crouching; good eyesight; intense concentration; ability to deal well with others; and work with computers. As we will discuss in the final section, these questions may shed light on issues concerning the vocational factors in SSA’s disability determination process. In addition, respondents are asked about the norms of their job (retiring at a certain age) or their desire for job changes as they age (fewer hours for the same hourly wage).

- **Workplace characteristics and accommodations.** Several studies have considered the role of workplace accommodations for workers with disabilities. One recent study by Hill
et al. (2014) found that only one in four older workers who experience disability receive an accommodation from their employer, and that the likelihood of doing so may be tied to individual personality attributes such as productivity and assertiveness. Work using earlier rounds of the HRS considered employer and individual responses to new work limitations (Crimmins and Hayward 2004; Daly and Bound 1996) and the impact of employer accommodations on workers seeking SSDI benefits (Burkhauser et al. 1999). Other studies have considered the effects of the Americans with Disabilities Act (Chikiros 2000; Charles 2004) and other state anti-discrimination laws (Burkhauser et al. 2011, Burkhauser et al. 2012).

- **Other services and supports.** The ability for older workers with health limitations to continue to work in their career job or in another employment arrangement may depend critically on the services and supports they are able to access. The HRS contains a range of questions that could be valuable to a researcher interested in these types of support, including the availability of a vehicle, public services and supports received, and the reliance on paid and unpaid caregivers. In addition, with proper permissions, researchers can access zip code, which would allow for the linking of respondents to local-level infrastructure information such as access to public transportation, as well as local labor market conditions.

**IV. Identifying the Receipt of Federal Disability Benefits in the HRS**

One of the key strengths of the HRS to study disabilities among older workers is the detailed information available in the survey about the application for and receipt of federal disability benefits from SSDI and SSI along with the receipt of SSR benefits beginning as early as the ERA. These data are available using respondent self-reports as well as using linked SSA administrative data; we discuss these options below.

Prior to turning to measurement of disability program participation, one note about the HRS sample size. As many researchers who have conducted work in this area with the HRS have discovered, sample sizes become quite small quickly when considering those who receive federal disability benefits. Approximately 10 percent of the U.S. population ages 50-64 receives disability benefits, meaning that among the younger cohorts sampled in the HRS (War Babies, Early Baby Boomers, and middle Baby Boomers), about 500 respondents will report SSDI receipt prior to retirement. Additional sample restrictions—particularly as researchers try to follow respondents over multiple waves—can make subgroup analyses prohibitively small. Although important findings can be generated from these samples, reductions in the size of future cohorts could significantly limit the HRS for research related to SSDI take-up.

**A. Self-reported SSDI and SSI application and claiming**
Self-reported information on SSDI and SSI application and receipt has been the main source of information used in HRS-based research on participation in those programs (Bhattacharya and Schoenbaum 2002; Borsch-Supan 2010; Bound and Waidmann 2010; Burkhauser et al. 2002; Giertz and Kubik 2011; Gruber and Kubik 2003; Li and Maestas 2008; McGarry and Skinner 2009; Schimmel and Stapleton 2012; Bound et al. 2013, Nicholas 2014). In each wave of the HRS, respondents are asked about whether they applied for SSDI or SSI, and if so what the outcome of that application was. If the respondent indicates a benefit award, additional questions ask about when benefits were first received, and if they were stopped. If the applicant was denied benefits, additional questions assess when the denial occurred and about reapplication. Linking respondent reports over pre-retirement waves of the HRS allows a researcher to trace out the application and claiming of SSDI and SSI from age 50 through FRA using publicly available HRS data, the same as any other data in the HRS. In theory, additional information from the earliest interview can offer information prior to age 50, but the reliability of that data, particularly regard to the timing of application and receipt might be questionable.

There are several issues that limit the self-reported data for research questions seeking detailed information about benefits. First, prior to wave 5 (survey year 2000), individuals were asked about SSDI and SSI together, meaning that it is impossible to discern to which program a person applied to or received benefits from based on the data as collected. One could presumably use information on work history or impute backward based on later reports of benefit receipt to make a reasonable assumption, but not without an unknown amount of error. Because wave 5 was about 8 years after the first interview of the HRS cohort, this means that much of the pre-FRA information for that cohort may be of limited value for certain research topics. Second, other cross-wave differences in how information was pre-loaded from earlier wave reports and how skip logic was implemented may affect the cross-wave comparability of certain measures (Chien et al. 2015). Finally, even in the absence of issues related to survey questions, respondents may misreport this information. Even in years when respondents are asked separately about SSDI and SSI, many applicants may be confused about which program they applied to. This may be especially important between ages 62 and FRA, when older workers with disabilities might seek early retirement, only to have an SSA representative suggest applying for SSDI instead. The extent to which respondents correctly report this, especially if they receive a retroactive benefit award, is unknown.

B. Administrative reports of SSDI and SSI application and receipt

Fortunately for those interested in a detailed assessment of application for and receipt of federal disability benefits, and interactions between those programs and SSR, the HRS offers a highly valuable linkage to SSA administrative records (Olson 1999). With proper permissions and funding (discussed below), users are able to access:
• **Cross-year benefit files**, which contain monthly Social Security information about program status and benefit payment amounts for SSDI and SSR, derived from SSA’s Master Beneficiary Record.

• **SSI files**, which contain monthly information about SSI program participation, derived from SSA’s Supplemental Security Record (SSR).

• **831 files**, which contain information about applications for SSDI and SSI and the outcomes of those applications.

• **Detailed and summary earnings files**, which contain historical information on earnings that are used in SSA’s computation of disability and retirement benefits.

Using the administrative records can avoid many of the pitfalls identified in using respondent reports, because the dates of application and award and monthly payment amounts are pulled directly from SSA’s reports. Of course, administrative data is not designed for research purposes, and thus, even these data may be limited in certain ways. For example, elements in SSA’s MBR are overwritten as benefits status changes (say, for example, when moving from SSDI to SSR), meaning that what a respondent reports at a certain point in time may not be what is contained in SSA’s administrative file at the time of the most recent pull from the HRS. In addition, because the files are pulled directly from administrative sources, they are missing important variables that have been constructed from SSA administrative data to support research as part of its Disability Analysis File (DAF).3 Perhaps the most important of these are a monthly indicator of non-payment status following suspense or termination due to work and a second variable indicating the benefits forgone due to work. The number of HRS respondents with matched data giving up benefits for work might be quite small, but seems at least to be worth checking.

Yet, despite the wealth of information, very few studies have taken advantage of these linked data sources. In our review of the literature, we identified only two completed studies that use the linked Cross-Year Benefit file (Bound et al. 2013 and Yin 2015). Several others have used the self-report of benefit receipt in combination with the summary earnings records, often as a way to overcome the inability to distinguish between SSDI and SSI in the earlier years of the study and relying on past earnings to determine eligibility for SSDI (Mitchell and Phillips 2001, Mitchell and Phillips 2000; Burkhauser et al. 2002). Based on our own experience accessing these data for two studies in progress, there are several reasons that researchers have not more fully utilized these data, most of which could be alleviated with better documentation:

1. **Data access procedures.** Accessing the SSA administrative records requires having a federally funded research project, completing a somewhat lengthy application, and securing IRB approval. By reputation, this process is onerous and can take many

3 SSA builds the DAF by aggregating information across more than a half dozen administrative data files, including those used to develop the data products offered through the HRS.
months. Our own experience suggests that this process can actually be relatively expeditious, with data access in a matter of weeks. Yet, those without federal funding are precluded; Bhattacharya and Schoenbaum (2002) noted: “However, access to these data is restricted and requires a federally-funded research project (among other criteria); the restricted data were thus not available for this study.”

2. Data documentation. Despite the restrictions for accessing the data, many researchers have been able to secure access, though primarily to work with the summary and detailed earnings records. Our own experience suggests that fewer researchers have attempted to access the administrative files related to benefit claiming because they are incredibly complex and the documentation available on the HRS site is sparse. The documentation that is available presumes a detailed understanding of program rules and file structures. It offers little in the way of user-friendly advice about getting started with the data or how to identify the most critical variables for certain research purposes. By contrast, the National Center of Health Statistics (NCHS) offers a similar linkage to SSA administrative files and has much more detailed documentation, often with a page description of each variable in the file and tips for using those variables. We consulted that documentation as much (or more) than the HRS documentation when working with the restricted files for the first time.

3. Inability to assess data quality *ex ante*. Faced with the process to seek approval to work with the files and limited documentation, many users would likely accept the limitations of the self-reported data. This is compounded by the fact that there is nothing immediately obvious on the HRS site that provides an assessment of data quality—the share of records that have a match in the administrative files, the share of respondents whose benefits are found to align with self-reported data. On the one hand, if researchers knew, for example, that 75 percent of respondents between ERA and FRA correctly reported their SSDI benefits or that 20 percent reported SSI when they actually had SSDI, they would be able to assess whether the value of the restricted data would make sense for their research purpose. On the other hand, if 95 percent of respondents report their benefits status correctly, many researchers could forego the restricted data. To our knowledge, no such information is available. Until late in 2015, there was a file that allowed users to determine the last year for which a respondent’s data was matched to SSA records, but the most useful variables in that file were stripped per SSA security requirements, and users were asked to destroy the file.

In addition to improving documentation, steps should be taken to ensure that HRS respondents continue to be asked to provide prospective permission to link their data. Prior to

---

4 Note that these numbers are completely fabricated for illustrative purposes.
2006, respondents provided permission to link their data through the interview date, but not beyond. This required asking beneficiaries to link their data again; our assessment of the data suggests that a non-trivial share either were not re-asked or did not agree to re-link their data. As a result, many in earlier cohorts of the HRS have linkage permission that ends prior to reaching FRA, limiting the value of the linked data for studying transitions prior to retirement. Since 2006, respondents have provided prospective permissions, only requiring their consent once to continue linking indefinitely into the future, significantly increasing the value of these data.

V. Conclusions

Information collected by the HRS over the past two decades has led to advances in the state of knowledge surrounding the consequences of disability onset (and relatedly, the onset of new health conditions) as well as the role of disability insurance and Social Security retirement benefits in mitigating its effects. This research has been made possible by a comprehensive and complex survey design to collect a wide range of information from respondents and spouses in a consistent fashion for many years before and after FRA. HRS innovations such as disability vignettes have pushed to the forefront of international disability research and have allowed new ways to assess work ability in the United States.

The breadth of information collected about health and functioning has allowed researchers to customize measures of disability and health to their own needs, often related to the age of the study population. Yet, the HRS surprisingly has not aligned its disability measures to other nationally representative surveys using the 6QS. While we do not suggest abandoning the existing measures, we do advocate for pilot testing the 6QS for at least a sample of HRS respondents in future years through an experimental module or another mechanism. This would allow researchers to benchmark data in the HRS to surveys such as the CPS and SIPP, better understand the extent to which existing measures yield a similar population of adults with disabilities, and better understand why substantial numbers of SSDI and SSI respondents in other surveys are missed by the 6QS disability measure.

The linkage to SSA administrative data expands the types of research questions that the HRS may be able to answer, but seems underutilized. Having definitive information about application for benefits, program participation, and monthly benefit amounts allows researchers the ability to better understand the interplay between SSDI, SSI, and SSR and household decisions surrounding seeking benefits. Maintaining prospective permissions for new cohorts will improve the likelihood that the matched sample has sufficient size and is representative of the HRS sample overall. HRS staff should also work to develop better documentation to support research SSA administrative files that can be linked to the core survey. The currently available resources require such detailed knowledge of SSA administrative data that those lacking that information would likely not have any idea how to use the linked files, nor understand their
potential value. A study that compares self-reported benefit information to information from the matched files and development of special weights for the matched records would make it easier for researchers to determine whether to use the matched data and help insure greater consistency in use across studies.5 Some consideration should also be given to adding administrative information about beneficiary return to work.

The HRS is positioned to help researchers answer questions that will be at the forefront of disability policy research for the indefinite future. For example, researchers are already using the HRS to explore the impact of the ACA Medicaid expansions on employment, and additional work might consider its impact on applications to SSI and SSDI using the linked administrative files. Future research will build on existing HRS-based research that has also already made substantial contributions to the evidence base needed to support substantial reforms to the SSDI, SSI and SSR programs that would have substantial consequences for older workers with significant health problems. Substantial Social Security and other policy changes that will affect this population seem inevitable as lawmakers grapple with projected funding shortfalls for SSR, SSDI and Medicare, and the overarching issue of growing federal deficits. There have been many proposals to accelerate the increase in the FRA, but they have been resisted at least in substantial part because of the consequences for older workers with health problems. Some policymakers have suggested phasing down SSDI benefit awards with age, to reduce or eliminate the disparity between SSDI and SSR benefits in the period between ERA and FRA. Others have suggested overhauling the use of vocational factors—age, education and job experience—in SSA’s disability determination criteria, or of increasing the age at which they start to apply. A large majority of SSDI entrants age 50 or older receive awards based on their residual functional capacity in combination with their age, education and past work, and analysis of the occupational, health and functional information in the HRS will likely contribute to the evidence needed to support changes to the use of the vocational factors. Maintenance and improvements in the HRS disability are critical for all of this important disability research, and more.

5 Our understanding is that a file with weights that account for matches to the summary earnings files provided through the SSA linkage are available, but that these do not apply to the other SSA administrative data products.
References


