Is the 2010 Affordable Care Act minimum standard to identify disability in all national datasets good enough for policy purposes?

Richard V. Burkhauser[a,*], T. Lynn Fisher[b], Andrew J. Houtenville[c] and
Jennifer R. Tennant[d]
[a]*Cornell University and University of Melbourne, Ithaca, NY, USA*
[b]*Social Security Administration, Woodlawn, MD, USA*
[c]*University of New Hampshire, Durham, NH, USA*
[d]*Ithaca College, Ithaca, NY, USA*

We provide a face validity test of the Department of Health and Human Services' (HHS) approved minimum six-question sequence (6QS) to capture the population with disabilities. Using linked 2009 Current Population Survey/Social Security Administration records data we find that the 6QS captures 66.3% of those whom administrative records confirm are receiving disability-based Social Security benefits. Adding a work-activity question increases our capture rate to 89.3%.

We find little difference in the distribution of conditions of those reporting only a 6QS-based disability and those only reporting a work activity-based disability. The four function-related questions do a relatively good job of capturing beneficiaries based on these conditions. But the work-activity question does a far better job of capturing beneficiaries than do the two activity-related questions. We conclude that the 6QS is fundamentally flawed and that any minimum standard for capturing the population with disabilities must include a work-activity question.

Keywords: Measuring disability populations, function-based measures, activity-based measures, disability prevalence, matched CPS/SSA administrative records data

1. Introduction

The Americans with Disabilities Act of 1990 was the culmination of a decades-long effort to afford people with disabilities the same protections against discrimination that the Civil Rights Act of 1964 or subsequent civil rights legislation provided those facing discrimination based on race, national origin, sex, and age. Unlike these other protected classes, whose characteristics are immutable or relatively easy to determine, disability is not an immutable characteristic. Hence, it is a more difficult characteristic to conceptualize and operationalize for purposes of affording civil

*Corresponding author: Richard V. Burkhauser, 259 MVR Hall, Cornell University, Ithaca, NY 14853, USA. Tel.: +1 607 255 2097; Fax: +1 607 255 4071; E-mail: rvb1@cornell.edu.

rights protection and, more generally, for statistical purposes including the monitoring of the health, employment, and economic well-being of this protected class. (See [11,13] for earlier reviews of the conceptual and methodological issues in measurement of work disabilities. See [1,8,10,16] for more recent reviews of disability measurement issues.)

To provide the information for evidence-based public policymaking it is necessary for researchers both inside and outside government agencies to have sufficient data to capture the effect of current and future policies on the classes of citizens that, based on past discrimination or current circumstances, require targeted government actions. With respect to people with disabilities this has meant efforts by government statistical agencies to develop a set of questions that could, within more general national datasets, identify the population with disabilities in a manner similar to the questions used to identify people by race, ethnicity, gender, and age.

In what is considered a milestone in the effort to determine a minimum standard for the set of questions required to identify disability in all national datasets, Section 4302 of the Affordable Care Act (ACA) of 2010 mandates the establishment of standards for the collection and dissemination of health statistics for five specific demographic sub-populations: race, ethnicity, gender, primary language, and disability status.

In response to this mandate, in October 2011, the U.S. Department of Health and Human Services (HHS) established new minimum data collection standards for these sub-populations, in which the new six-question sequence on disability (6QS) that was introduced in the American Community Survey (ACS) and is currently used in the Current Population Survey (CPS), was deemed "the data standard for survey questions on disability" [9]. The first four questions of the 6QS are function-based, focusing on: hearing, vision, cognition (concentrating, remembering, or making decisions), and mobility (physical matters like walking and climbing stairs). The last two are activity-based, relating to activities of daily living (dressing or bathing) and instrumental activities of daily living (doing errands alone such as visiting a doctor's office or shopping). See Appendix Table 1 for their specific wording and that of other variables.

The American Housing Survey (AHS), National Health Interview Survey (NHIS), the incoming rotations of the CPS-Basic Monthly Survey (CPS-BMS), and the American Time Use Survey (ATUS) include the 6QS. In their report, HHS cites the use of the 6QS in national surveys and the Office of Management and Budget's encouragement of other federal agencies to use the 6QS because of the "extensive testing used in the development of these measures, including the findings that alternative measures did not test as well" as part of their justification for making it "the data standard for survey questions on disability" [9].

Using Social Security Administration (SSA) administrative records data, we provide evidence that the lack of a work-activity question in the 6QS results in its inability to capture a substantial portion of the population with disabilities that should

be included in any working-age population with disabilities – those currently receiving Social Security benefits based on their disability. Since this Type 2 error only comes from a population of current program recipients who do not answer "yes" to the 6QS, our approach understates the number of Type 2 errors. We are unable to capture Type 2 error arising from those who do not answer "yes" to the 6QS, or who have disabilities but are not currently receiving either Social Security Disability Insurance (SSDI) or Supplemental Security Insurance – Disabled Adult (SSI-Disabled Adult) benefits.[1]

Our face validity test uses linked 2009 Current Population Survey-Annual Social and Economic Supplement (CPS-ASEC)/SSA administrative records data. We find that the 6QS captures only 66.3% of those who administrative records confirm are receiving Social Security benefits based on their disability. When we add a work-activity question to the 6QS our capture rate increases by another 23.1 percentage points – for a total of 89.3%.

Furthermore, we find little difference in the distribution of conditions between those who only self-report a 6QS-based disability and those who only self-report a work activity-based disability. Each of the four function-related questions in the 6QS do a relatively good job of capturing those receiving Social Security benefits based on a condition the literature suggests is related to that function. But the work-activity question does a far better job of capturing those receiving Social Security benefits than the two activity-related questions in the 6QS.

In 2006, Sallie Keller-McNulty, the then-President of the American Statistical Association, urged that research on technical and methodological adjustments to a work-activity question continue so that it could be added to the ACS to improve the measurement of work disability [12]. We find that a work-activity question in combination with the 6QS substantially improves the ability to identify a disability subpopulation that should be included in any population of working-age people with disabilities – those receiving Social Security benefits based on their disability. We conclude that the absence of a work-activity question in the 6QS makes it a fundamentally flawed data standard for survey questions on disability.

2. Concepts of disability

There is no universal agreement on the most appropriate definition of disability, although the World Health Organization's International Classification of Disability,

---

[1] We do not address Type 1 error (false positives) because there is not a convenient sample of individuals known to not have a disability, since individuals not receiving SSDI/SSI-Disability income may or may not have a disability. Even those whose SSDI/SSI-Disability applications have been rejected cannot be known to not have a disability. Past denial of benefits is not a suitable method of identifying absence of disability.

Health and Functioning (ICF) [18] is a commonly used framework for defining disability. In the ICF framework, the term disability describes the health condition-based presence of an impairment, activity limitation, and/or participation restriction. The emergence of the ICF as a systematic and comprehensive way of conceptualizing the population with disabilities has resulted in an international effort to use these classifications to better identify the population with disabilities in government-sponsored datasets [17].

A major challenge is how to operationally identify a random sample of this complex conceptualization of the population with disabilities in a national survey where questionnaire space is highly constrained. One way to describe how the CPS-ASEC may be used to do so is to imagine a square containing the entire population with health conditions. Within the square are three concentric circles (i.e., in the shape of an archery target), with the outermost circle representing people with disabilities using ICF concepts (that is, having health condition-based impairments, activity limitations, or participation restrictions), the middle circle representing those with work-activity limitations (a subset of the broader ICF-defined population), and the innermost circle representing people currently receiving Social Security benefits based on their disability (a subset of the work-activity limited population whose limitations are severe enough to prevent them from performing "any substantial gainful activity" – that is, a sub-population whose work-activity limitations are severe enough to meet the eligibility criteria for these permanent and total Social Security disability programs).

We will show that although the outermost circle is the concept of disability that the 6QS is attempting to operationally achieve, the 6QS-based and work-activity-based disability populations are best described by a Venn diagram with the majority of people either responding "yes" to one of the 6QS questions and "no" to the work-activity question or vice versa, and only a minority responding "yes" to both. More importantly, as we will discuss in more detail, when we use SSA administrative records linked to the CPS data to determine more precisely who is currently receiving Social Security income based on their disability (as opposed to the self-reporting of disability benefits), our Venn diagram of that population shows that only 66.3% of these beneficiaries who should be included in any conceptualization of the disability population are captured by the 6QS alone.

3. Data and key variables

Burkhauser et al. [2] were the first to use data from the public-use CPS to show how sensitive the size and socioeconomic characteristics of the working-age population with disabilities are to the questions used to capture that population. They proposed a face validity test of the relative merits of each. That is: How well do the questions capture a subpopulation that should be included in any formulation of

working-age people with disabilities – current SSDI and SSI-Disabled Adult income recipients?

However, they based their conclusions on results using the 2010 public-use CPS data that only contains self-reported information on receipt of Social Security income. Here we use two datasets: (1) the 2009 public-use CPS-ASEC and (2) 2009 public-use CPS-ASEC data matched to SSA administrative records, which was the most recent matched data available. These matched data provide much more accurate information on who in the CPS data is currently receiving Social Security benefits based on their disability and the main diagnosis (medical listing) on which that disability is based.

3.1. General information about the CPS

The CPS is a joint undertaking of the Bureau of Labor Statistics (BLS) and the Census Bureau. It is a monthly survey of approximately 57,000 households and is the primary source of labor force information on the working-age (16 years and over), non-institutionalized population. The CPS-BMS data contain labor force information and demographic information, and since June 2008 include a series of six disability questions – see Appendix Table 1. In June 2008, the 6QS was asked of all respondents in that month-in-sample. Thereafter it was asked in the respondents' first and fifth months-in-sample [5]. As can be seen in Appendix Table 1, the first four questions are function-based, focusing on: hearing, vision, cognition (concentrating, remembering, or making decisions), and mobility (physical matters like walking and climbing stairs). The last two are activity-based, relating to activities of daily living (dressing or bathing) and instrumental activities of daily living (doing errands alone such as visiting a doctor's office or shopping).

Prior to 2008, the ACS included a work-activity question as part of its original 6QS. (See Appendix Table 1A for the wording of this question.) But this question was dropped in 2008. The CPS-BMS 6QS – patterned after the 2008 ACS – also does not include a work-activity question, although a work-activity question has been asked in the CPS-ASEC as part of this supplement to the CPS-BMS each March since 1981. (See Appendix Table 1.) The CPS-ASEC provides the usual monthly labor force data provided in the BMS, but also adds data on work experience, income, non-cash benefits, and migration [4]. It also includes and separates the many sources of household income, including Social Security benefits.

The CPS uses a rotation system for its interviews. Each housing unit is followed for a 16-month period – four months in-sample, eight months out-of-sample, and then four months in-sample. That is, a respondent in a selected housing unit is interviewed with respect to all persons living in that housing unit for four consecutive months. After eight consecutive months without being interviewed, a respondent in that housing unit is interviewed for another four consecutive months, after which the housing unit is retired from the CPS sample. In any sample month, one-eighth of

the sample is being interviewed for the first time (month-in-sample, or MIS = 1), one-eighth is being interviewed for the second time (MIS = 2), and so on [3].

Because of the 4-8-4 rotation system described above, in the absence of any matching issues, a person who is in his first month-in-sample would be interviewed one year (12 months) later but would only be in his fifth month-in-sample.[2] The 2009 March CPS-ASEC was administered after the 6QS was in effect, so in that dataset, everyone is asked the work-activity limitation question in March, and the 6QS at some point before, or concurrently if March is the first or fifth month-in-sample for that particular household. So, the March 2009 CPS-ASEC dataset gives us the information we need to compare these two measures of disability.

3.2. Match process of the restricted use file

For this study, we matched the 2009 CPS-ASEC public use file to the Social Security Account Number Identification (NUMIDENT), Master Beneficiary Record (MBR), and Supplemental Security Record (SSR) files.[3] The NUMIDENT file contains information on all persons who have ever submitted an application for a Social Security Number; the MBR file contains the records of the Old-Age, Survivors, and Disability Insurance program; the SSR file contains the records of the SSI program [7]. The statistics in this study were generated using SAS 9.3 (TS1M2), Rev. 930_12w41 on an Intel Core2 Duo CPU E8400 running at 3.00 GHz with 4 GB of memory, running Windows 7 Enterprise, Service Pack 1, 32-bit operating system.

3.3. Match rate

SSA records were matched for 90,001 of the 102,726 persons ages 25–61 in the 2009 CPS-ASEC sample – a match rate of 87.6%.[4] This means that 12,725 (12.5% of) persons ages 25–61 in the 2009 CPS-ASEC sample were not matched to SSA records. From a population perspective, using the CPS-ASEC sample weights, SSA records were matched for 131,881,301 of the 152,003,928 non-institutionalized working-age (aged 25–61) United States population – a weighted match rate of 86.8%.

---

[2]MIS = 2 and MIS = 6, MIS = 3 and MIS = 7, and MIS = 4 and MIS = 8 also create yearlong matched samples.

[3]The matched data are available on a restricted basis to researchers with special sworn status from the Census Bureau, working on approved projects at restricted data sites, subject to the terms of an interagency agreement between the Census Bureau and SSA. SSA's Title 13 disclosure review board approved all our estimates prior to their distribution.

[4]The match rate is less than 100% for several reasons, including respondents not consenting to the match and the inability to determine the Social Security numbers of respondents who consented to the match.

3.4. Sample weights

We rescaled our sample population using the CPS-ASEC sample weights to reestablish population representativeness of matched records across 16 age-gender subpopulations (gender by eight age groups: 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, and 60–61). We did so by multiplying an individual's sample weight by their rescaling factor for his/her age-gender subpopulation, where the rescaling factor is equal to the ratio of the subpopulation size of the full sample to the subpopulation size of the matched sample as recommended by Czajka, Mabli, and Cody [6]. Here, we only report results for our weighted matched sample. Tables using the full CPS sample are available in an unpublished appendix available from the authors. The results are similar.

3.5. Defining receipt of disability income

Matching CPS-ASEC and SSA administrative records is a more precise way to determine if an individual is receiving Social Security income based on a disability. SSDI income receipt is based on whether an individual was (a) in the MBR file, (b) received benefits based on a disability (as opposed to retirement or survivorship), (c) had a current payment, and (d) is eligible based on his/her own contributions or the contributions of his/her retired parent(s) or deceased spouse. This last inclusion criterion means that in addition to Disabled Workers, individuals receiving benefits under the Disabled Adult-Child (DAC) program or the Disabled Widows/Widowers program are also included, providing that the other criteria are also met.

Receipt of SSI-Disabled Adult income is based on whether an individual: (a) was in the SSR file, (b) was under age 65, and (c) has a current payment.

4. Results

When we use our 2009 matched-CPS data, we do not get the concentric circle model suggested by theory. The work-activity-based disability population (B + C) in the Venn diagram pictured in Fig. 1 is not a subset of the 6QS-based disability population (A + B), but only partially overlaps this group. Only when the population that forms the union of these two populations – the seven-question (7Q)-based disability population (A + B + C) – is considered to be the outermost circle are we able to successfully operationalize our concentric circle conceptualization of disability. However, only about 40% of the people in this broader disability category are the same people (that is, have positively answered both a question in the 6QS and the work-activity question – B in the diagram). Of the rest, 29% have answered positively to the 6QS but not the work-activity question (A), and 31% have responded positively to the work-activity question but not the 6QS (C). Hence using either the
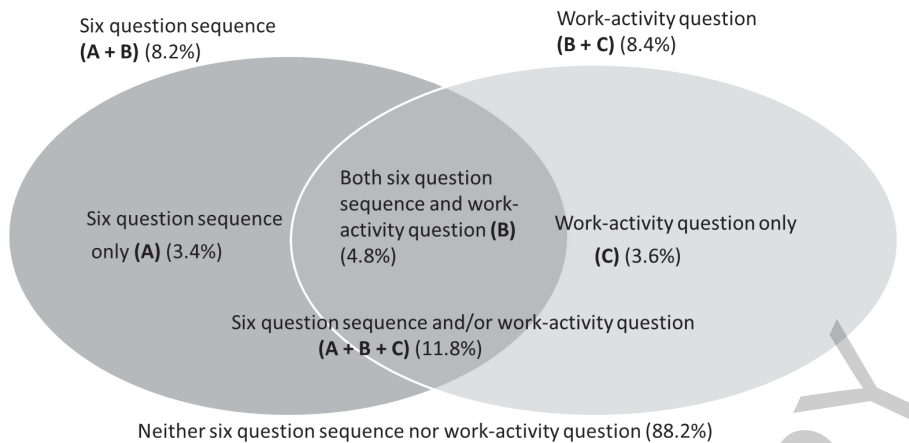
Six question sequence
(A + B) (8.2%)

Work-activity question
(B + C) (8.4%)

Six question sequence
only **(A)** (3.4%)

Both six question
sequence and work-
activity question **(B)**
(4.8%)

Work-activity question only
(C) (3.6%)

Six question sequence and/or work-activity question
(A + B + C) (11.8%)

Neither six question sequence nor work-activity question (88.2%)

Fig. 1. Prevalence rate of non-institutionalized civilians ages 25–61, by type of disability population measure, 2009 Matched SSA/CPS-ASEC.

6QS or the work-activity-based question alone to capture the population with disabilities will dramatically understate the population formed by using all seven questions (A + B + C).

Rows 1 and 2 of Table 1 put the magnitudes of the populations in the Venn diagram into perspective. Of the 152 million Americans aged 25-61 in our weighted sample, 11.8% or 17.9 million have disabilities based on our 7Q-based definition (A + B + C). These values fall to 8.2% or 12.5 million when using solely the 6QS (A + B). Hence using the 6QS, which is the HHS-established minimum data collection standard for people with disabilities, will miss the 3.6% or 5.5 million working-age people who self-report having a work-activity limitation but say no to each of the six questions in the 6QS (C).

Table 1 also shows that while the overall 6QS-based disability prevalence is about the same as the work-activity-based disability prevalence, there are dramatic differences in these two populations' labor force participation, employment, and poverty rates. Those in the 6QS population (A + B) are much more likely to be in the labor force (36.8 vs. 22.3%) or employed (32.0 vs. 17.8%) and less likely to be in poverty (24.8 vs. 28.8%) than those in the work-activity population (B + C). More importantly, the 31% of the broader 7Q-based disability population who are in the work-activity-only disability subpopulation (C) have dramatically lower labor force (31.9 vs. 68.5%) and employment (26.8 vs. 60.5%) and dramatically higher poverty (26.8 vs. 16.7%) rates than the 29% who are in the 6QS-only disability subpopulation (A).

These differences in economic outcomes are accounted for to some degree in Row 6 which uses our administrative records data to show the dramatic differences in the share of those disability populations who are currently receiving SSDI income or SSI-Disabled Adult benefits. Part of the reason for this difference is that the severity

Table 1

Population size, disability prevalence rate, and economic outcomes of non-institutionalized persons ages 25–61, by type of disability population measure, using the 2009 Matched SSA/CPS-ASEC[a]

Variable	Total population	Six question sequence (A+B)	Work activity (B+C)	Six question sequence-only (A)	Both six question sequence and work activity (B)	Work activity-only (C)	Six question sequence and/or work activity (A+B+C)	Neither six question sequence nor work activity
Disability prevalence rate	–	8.2	8.4	3.4	4.8	3.6	11.8	88.2
Population size	152,000,000	12,500,000	12,800,000	5,200,000	7,300,000	5,500,000	17,900,000	134,100,000
Labor force participation rate	81.3	36.8	22.3	68.5	13.6	31.9	35.4	87.4
Employment rate	74.7	32.0	17.8	60.5	12.4	26.8	29.9	80.7
Poverty rate	9.3	24.8	28.8	16.7	30.3	26.8	25.4	7.1
Pct. DI and/or SSI[b]	6.3	51.5	57.7	21.0	73.4	38.4	47.3	0.8

[a]The 2009 CPS-ASEC file was matched to the SSA Numident file. Numident records were matched for 90,001 of the 102,726 non-institutionalized persons ages 25–61 in the 2009 CPS-ASEC sample – a match rate of 87.6%. Hence 12,725 or 12.5% of persons ages 25–61 in the 2009 CPS-ASEC sample were not matched to Numident records. The CPS-ASEC sample weight was rescaled based on gender and age to account for non-matching sample and reestablish population size estimates.

[b]SSDI status is based on whether an individual was (a) in the MBR file, (b) receiving benefits on the basis of disability (as opposed to retirement or survivorship), and (c) had a current payment. SSI status is based on whether an individual was (a) in the SSR file, (b) under age 65, and (c) has a current payment.

of the impairments of those who are included in each of the various ways of creating a disability population will vary. Note that in the disability subpopulation of persons who report a 6QS-based disability and a work-activity-based disability (B), 73.4% are receiving disability benefits. In contrast only 21.0% of those in the 6QS-only subpopulation (A) and 38.4% of those in the work-activity-only subpopulation (C) do so.

However, as shown below, these differences in economic outcomes are also in part accounted for by the ability of the questions used to create these various samples of the true disability population to actually capture those who administrative record data objectively tell us are receiving Social Security disability benefits.

Therefore, using either the 6QS (A + B) or work-activity (B + C) population will not only understate the larger population with disabilities captured by the seven-question union of these populations (A + B + C), but is likely to create biased estimates of the labor force, employment, and poverty rates of this broader population with disabilities in part because of these questions' different abilities to capture those currently receiving SSDI or SSI-Disabled Adults benefits.

4.1. A face validity test

Prior to 2008, the American Community Survey (ACS) included a similar work-activity question that is still used in the CPS-ASEC as part of its original 6QS (See Appendix Tables 1 and 1A). In 2008, the ACS disability questions were substantially revised, splitting hearing and vision into separate questions and removing its work-activity question.

The scientific evidence for using the 6QS contained in the revised ACS was based on cognitive testing of how well respondents understood the questions and provided accurate answers. Conducted by an interagency committee using a small non-representative sample, these tests consisted of 69 interviews held over five rounds (an average of 14 interviews per round) in which respondents were probed with respect to their understanding of the questions and the basis for their responses. In each round, the questions were revised to improve respondent understanding. The final results of this testing were judged based on whether the survey respondents' answers to the questions demonstrated understanding in a manner sufficiently consistent with the Committee's expectations. The decision to remove the work-activity question in the ACS was made in the third round [14].

In an April 20, 2006, letter to the Census Bureau Director C. Louis Kincannon, Keller-McNulty [12], the then-President of the American Statistical Association, urged that research on technical and methodological adjustments to a work-activity question continue so that it could be added to the ACS to improve the measurement of work disability. This advice was not followed and the work-activity question was dropped from the 6QS in the ACS in 2008. This same 6QS without a work-activity limitation question was also included in the CPS-BMS starting in June 2008. It is now the HHS "data standard for survey questions on disability."

Below we provide a face validity test of the suitability of excluding a work-activity question in a dataset that is meant to capture the population with disabilities who are targeted for government services. The SSA runs two programs that provide income to working-age people based on their disability. The SSDI program provides Disabled Worker benefits, based on their past earnings, to workers who administrators determine to be disabled – those unable to perform substantial gainful activity because of their disability. The SSDI program also provides a DAC benefit to unmarried adults (aged 18 or older) who administrators determine to be currently disabled, based on these same SSDI disability criteria, if the disability began before age 22 and they have a parent who is deceased or is currently receiving retirement or disability benefits. It is called a child's benefit because it is paid based on the parent's Social Security earning record. SSDI benefits may also be paid to a Disabled Widow(er) with disabilities (or surviving divorced spouse with a disability), if he or she is between ages 50 and 60, their condition meets the definition of disability for adults, and the disability started before or within seven years of the worker's death.

The second program, SSI-Disabled Adult, provides a means-tested guaranteed benefit to those who administrators determine to be disabled based on the same SSDI disability criteria.

These programs are targeted to working-age people with disabilities who are unable to earn a minimum amount of income based on an impairment stemming from their health condition. The severity of the work-activity limitation required to enter these programs is obviously within the ICF conceptualization of disability. Most importantly with respect to the representativeness of the population captured by our alternative measures of the disability population, as a class the sub-population currently receiving these benefits is less likely to be either in the labor force or employed and more likely to be in poverty. Hence one face validity test of any sequence of questions used to capture the entire disability population and its economic characteristics is its ability to capture this part of the disability population.

When we use our 2009 matched-CPS data in Fig. 2, the population receiving Social Security benefits (SSDI and SSI-Disabled Adult) based on their responses to the 2009 matched-CPS are not a subsample of those self-reporting a work-activity limitation or even of those reporting one of the 6QS-based disability questions. Figure 2 is a Venn diagram that divides the 9.5 million adults (aged 25-61) using the same self-reported disability subsets described in Fig. 1. The 6QS (A + B) is able to capture 66.3% of this population, missing the 23.1% who only report a work-activity (C). Likewise, while the work-activity-based population question (B + C) captures 78.0% of this population, it misses the 11.3% of beneficiaries who only report one of the six impairment/activity limitations questions (A). Together, the union of these questions (A + B + C) captures 89.3% of the adult Social Security population receiving benefits based on their own disability.

Despite its greater number of questions, the 6QS alone is less able to identify our administrative-record-defined sample of SSDI and SSI beneficiaries than the work-activity question alone. This suggests that the failure to include some form of work-
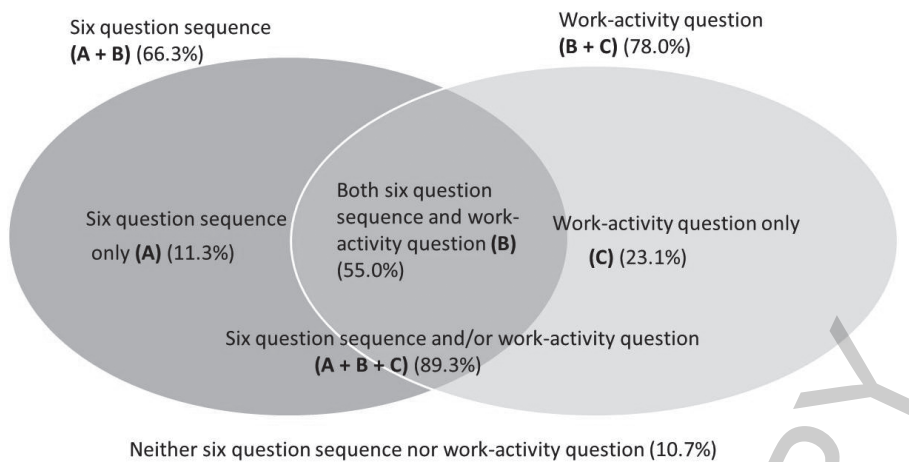
Fig. 2. Percentage of Social Security program beneficiaries captured by type of disability population measure, 2009 Matched SSA/CPS-ASEC.

activity question in a set of questions aiming to capture the broader disability population will substantially undercount the number of persons actually receiving SSDI or SSI-Disabled Adult benefits.

Because the labor force participation and employment rates of these missing beneficiaries are likely to be substantially lower than those of the rest of the working-age population with disabilities, their absence from the officially measured population with disabilities in the 6QS (A + B) is one reason for its likely upward bias in labor force participation and employment rates and its downward bias in poverty rates. What Fig. 2 demonstrates is that the HHS "data standard for survey questions on disability" substantially underestimates the prevalence of SSDI and SSI-Disabled Adults beneficiaries in the true disability population.

Row 1 of Table 2 repeats the percentages of the 9.5 million people in our weighted CPS data receiving SSDI and SSI-Disabled Adult benefits that are captured by each of our disability populations based on 6QS and work limitation self-reports. The next two rows of Table 2 do so for the 7.2 million who receive SSDI and the 3.4 million who receive SSI-Disabled Adult benefits. There is little difference in the capture rates of any of the disability populations between these two programs. The sample size for the DAC and Disabled Widow(ers) sample is too small to separate them from the recipients with previous work histories

Tables 3a and 3b take further advantage of the added information provided by the Social Security administrative records that are contained in the 2009 matched-CPS dataset. In addition to allowing us to capture the 9.5 million working-age persons who are receiving Social Security benefits based on their disability and how they are distributed across the same self-reported disability subsets described in Fig. 2, it also allows us to show how these distributions vary across the diagnosis groups of their primary medical condition.

Table 2
Social Security program beneficiaries[a] and percentage captured by type of disability population measure, using the 2009 Matched SSA/CPS-ASEC[b]

				Share captured by disability population measure				
Variable	Social Security population size	Six question sequence (A+B)	Work limitation (B+C)	Six question sequence-only (A)	Both six question sequence and work limitation (B)	Work activity-only (C)	Six question sequence and/or work limitation (A+B+C)	Neither six question sequence nor work limitation
DI and/or SSI	9,500,000	66.3	78.0	11.3	55.0	23.1	89.3	10.7
DI	7,200,000	66.9	79.2	10.9	56.0	23.2	90.1	9.9
SSI	3,400,000	65.0	75.2	12.7	52.3	22.9	87.9	12.1

[a] SSDI status is based on whether an individual was (a) in the MBR file, (b) receiving benefits on the basis of disability (as opposed to retirement or survivorship), and (c) had a current payment. SSI status is based on whether an individual was (a) in the SSR file, (b) under age 65, and (c) has a current payment.
[b] The 2009 CPS-ASEC file was matched to the SSA Numident file. Numident records were matched for 90,001 of the 102,726 non-institutionalized persons ages 25–61 in the 2009 CPS-ASEC sample – a match rate of 87.6%. Hence 12,725 or 12.5% of non-institutionalized persons ages 25-61 in the 2009 CPS-ASEC sample were not matched to Numident records. The CPS-ASEC sample weight was rescaled based on gender and age to account for non-matching sample and reestablish population size estimates.

We divide the data into the nine largest primary diagnosis groups and combine the rest into a tenth "other" diagnosis classification. These groups are reported from highest to lowest in terms of their populations in Table 3a and by their prevalence in Table 3b. The two largest groups – "Mental Disorders other than intellectual disabilities" and "Musculoskeletal" – account for 4.12 million or 43.6% of all beneficiaries. Since 15.1% of our diagnosis categories contain blank or invalid codes, these two largest diagnosis classifications account for 51.3% of all valid responses.

The first row of Table 3a shows how the 9.5 million working-age persons receiving Social Security benefits based on their disability are distributed by how they self-report their disability. It effectively reports the total number of persons in each of the Venn diagram categories reported in Fig. 2. The remaining rows show how this distribution changes across our self-reported CPS disability populations. The number of disability beneficiaries missed by using a 6QS is greater, the greater the diagnoses group size. Of the 2,240,000 with a primary diagnosis of "Mental Disorders other than intellectual disability," 540,000 fall into the work-activity-only subpopulation (C). Hence they are missed by the 6QS. The number missed among those whose primary diagnoses is "Musculoskeletal" is 440,000.

The first row of Table 3b shows the percentage of the 9.5 million working-age persons receiving benefits in each of the disability subpopulations used in the Venn diagram in Fig. 2. The remaining rows show how these percentages change in each diagnosis group. There is some variation by diagnosis group, but in all but one case, the work-activity-based disability population (B + C) captures a larger share than the 6QS-based disability population (A + B).

The only exception is "Nervous System and Sense Organs." In this case 82.7% are captured by the 6QS population and 80.2% are captured in the work-activity population. Note that two of the four function-based questions in the 6QS (hearing and vision) fall within the "Nervous System and Sense Organs" diagnosis group. This may explain why the 6QS does its best job, by far, of capturing beneficiaries in this diagnosis group. It is also the only diagnosis group where the addition of the 6QS-only population (A), 13.3%, to the work-activity population (B + C) exceeds the addition of the work activity-only population (C), 10.8%, to the 6QS-population (A + B).

Tables 3a and 3b reinforce the point that adding the work-activity question to the 6QS would substantially increase the share of current Social Security beneficiaries – based on administrative records data – captured with this broader seven-question measure of disability.

In Table 1 we observed wide variations in labor force participation, employment, and program participation rates across our disability populations. Thus including or excluding those who report a work-activity limitation only (C) importantly affected these key characteristics and hence our estimates of these key characteristics in the "true population."

Table 4 focuses on how the distribution of diagnoses varies across our disability populations based on self-reported 6QS and work-activity questions. It shows that

Table 3a
Social Security program beneficiaries[a] and percentage captured by type of disability population measure, using the 2009 Matched SSA/CPS-ASEC[b]

Diagnosis group	Social Security total	Six question sequence (A+B)	Work limitation (B+C)	Number captured by disability population measure				
				Six question sequence-only (A)	Both six question sequence and work limitation (B)	Work activity-only (C)	Six question sequence and/or work limitation (A+B+C)	Neither six question sequence nor work limitation
All groups	9,500,000	6,280,000	7,390,000	1,070,000	5,210,000	2,190,000	8,460,000	1,010,000
Mental disorders other than intellectual disability	2,240,000	1,500,000	1,800,000	260,000	1,200,000	540,000	2,000,000	230,000
Musculoskeletal system	1,880,000	1,200,000	1,500,000	200,000	1,000,000	440,000	1,700,000	200,000
Nervous system and sense organs	810,000	670,000	650,000	110,000	560,000	90,000	760,000	50,000
Intellectual disability	770,000	490,000	590,000	90,000	400,000	190,000	680,000	90,000
Circulatory system	620,000	400,000	480,000	70,000	330,000	150,000	550,000	70,000
Injuries	380,000	250,000	310,000	40,000	210,000	100,000	350,000	30,000
Endocrine, nutritional, metabolic	300,000	200,000	250,000	20,000	180,000	70,000	280,000	30,000
Respiratory system	200,000	120,000	170,000	20,000	100,000	70,000	190,000	10,000
Neoplasms	190,000	90,000	140,000	20,000	80,000	60,000	150,000	40,000
Other	640,000	420,000	480,000	72,000	350,000	143,000	560,000	83,000
Blank/invalid code	1,430,000	900,000	1,100,000	170,000	740,000	340,000	1,200,000	180,000

[a] SSDI status is based on whether an individual was (a) in the MBR file, (b) receiving benefits on the basis of disability (as opposed to retirement or survivorship), and (c) had a current payment. SSI status is based on whether an individual was (a) in the SSR file, (b) under age 65, and (c) has a current payment.
[b] The 2009 CPS-ASEC file was matched to the SSA Numident file. Numident records were matched for 90,001 of the 102,726 non-institutionalized persons ages 25-61 in the 2009 CPS-ASEC sample – a match rate of 87.6%. Hence 12,725 or 12.5% of non-institutionalized persons ages 25-61 in the 2009 CPS-ASEC sample were not matched to Numident records. The CPS-ASEC sample weight was rescaled based on gender and age to account for non-matching sample and reestablish population size estimates.

Table 3b
Social Security program beneficiaries[a] and the percent captured, by diagnosis group and type of disability population measure using the 2009 Matched SSA/CPS-ASEC[b]

Diagnosis group	Social Security total	Disability population measure						
		Six question sequence (A+B)	Work limitation (B+C)	Six question sequence-only (A)	Both six question sequence and work limitation (B)	Work activity-only (C)	Six question sequence and/or work limitation (A+B+C)	Neither six question sequence nor work limitation
All groups	9,500,000	66.3	78.0	11.3	55.0	23.1	89.3	10.7
Mental disorders other than intellectual disability	2,240,000	65.9	78.2	11.6	54.4	23.9	89.8	10.2
Musculoskeletal system	1,880,000	66.1	78.6	10.8	55.3	23.3	89.4	10.6
Nervous system and sense organs	810,000	82.7	80.2	13.3	69.4	10.8	93.5	6.5
Intellectual disability	770,000	64.5	76.9	11.8	52.6	24.3	88.8	11.2
Circulatory system	620,000	64.6	77.2	11.3	53.4	23.8	88.5	11.5
Injuries	380,000	66.6	82.3	10.3	56.3	26.0	92.6	7.4
Endocrine, nutritional, metabolic	300,000	66.6	83.0	8.1	58.5	24.6	91.1	8.9
Respiratory system	200,000	57.8	83.7	9.6	48.3	35.5	93.3	6.7
Neoplasms	190,000	49.1	70.9	7.9	41.2	29.6	78.7	21.3
Other	640,000	65.5	76.0	11.5	53.9	23.2	87.2	12.8
Blank/invalid code	1,430,000	63.4	75.4	11.7	51.6	23.8	87.2	12.8

[a]SSDI status is based on whether an individual was (a) in the MBR file, (b) receiving benefits on the basis of disability (as opposed to retirement or survivorship), and (c) had a current payment. SSI status is based on whether an individual was (a) in the SSR file, (b) under age 65, and (c) has a current payment.
[b]The 2009 CPS-ASEC file was matched to the SSA Numident file. Numident records were matched for 90,001 of the 102,726 non-institutionalized persons ages 25–61 in the 2009 CPS-ASEC sample – a match rate of 87.6%. Hence 12,725 or 12.5% of persons ages 25–61 in the 2009 CPS-ASEC sample were not matched to Numident records. The CPS-ASEC sample weight was rescaled based on gender and age to account for non-matching sample and reestablish population size estimates.

Table 4
Percentage of Social Security program beneficiaries[a] by diagnosis group and type of disability population measure using the 2009 Matched SSA/CPS-ASEC[b]

Diagnosis group	Social Security total	Six question sequence (A+B)	Work limitation (B+C)	Disability population measure				
				Six question sequence-only (A)	Both six question sequence and work limitation (B)	Work activity-only (C)	Six question sequence and/or work limitation (A+B+C)	Neither six question sequence nor work limitation
All groups	9,500,000	6,280,000	7,390,000	1,070,000	5,210,000	2,190,000	8,46,000	1,010,000
Mental disorders other than intellectual disability	23.7	23.5	23.7	24.3	23.4	24.5	23.8	22.5
Musculoskeletal system	19.9	19.8	20.0	19.0	20.0	20.0	19.9	19.7
Nervous system and sense organs	8.6	10.7	8.8	10.1	10.8	4.0	9.0	5.2
Intellectual disability	8.1	7.9	8.0	8.5	7.7	8.5	8.0	8.5
Circulatory system	6.6	6.4	6.5	6.6	6.4	6.8	6.5	7.1
Injuries	4.0	4.0	4.2	3.7	4.1	4.5	4.2	2.8
Endocrine, nutritional, metabolic	3.2	3.2	3.4	2.3	3.4	3.4	3.3	2.7
Respiratory system	2.2	1.9	2.3	1.8	1.9	3.3	2.3	1.3
Neoplasms	2.0	1.5	1.9	1.4	1.5	2.6	1.8	4.1
Other	6.6	6.7	6.6	6.7	6.7	6.9	6.5	8.1
Blank/invalid code	15.1	14.4	14.6	15.6	14.1	15.5	14.7	18.0

[a]SSDI status is based on whether an individual was (a) in the MBR file, (b) receiving benefits on the basis of disability (as opposed to retirement or survivorship), and (c) had a current payment. SSI status is based on whether an individual was (a) in the SSR file, (b) under age 65, and (c) has a current payment.

[b]The 2009 CPS-ASEC file was matched to the SSA Numident file. Numident records were matched for 90,001 of the 102,726 non-institutionalized persons ages 25–61 in the 2009 CPS-ASEC sample – a match rate of 87.6%. Hence 12,725 or 12.5% of persons ages 25–61 in the 2009 CPS-ASEC sample were not matched to Numident records. The CPS-ASEC sample weight was rescaled based on gender and age to account for non-matching sample and reestablish population size estimates.

the variation in this distribution (e.g. in Row 2, the share of "Mental Disorders other than intellectual disabilities" is very similar across all columns, etc.) is much less pronounced than it was for labor force participation, employment, or poverty rates in Table 1. Furthermore, the diagnoses distribution in the work-activity-only population (C) is not dramatically different from the diagnoses distribution in 6QS population (A + B) or even in the 6QS-only population (A).

The rank order of the share of conditions is quite close across all the disability populations in Table 4. In fact, the rank order of conditions in the 6QS (A + B) perfectly matches that of the work-activity population (B + C).

Even when we compare our 6QS-only (A) and work-activity-only (C) populations, the rank order is remarkably close. The only major anomaly in this comparison is that the "Nervous System and Sense Organ" diagnosis group is ranked third (10.1% share) in the former and sixth (4.0% share) in the latter. As discussed above, two of the four function-based questions in the 6QS (hearing and vision) fall within this diagnosis group. As we will discuss in more detail below, most of this difference is likely to be related to these subcategories.

Above we have argued that the 9.5 million SSDI and SSI persons in our CPS sample based on administrative records should be in any disability population. Tables 3a and 3b show that expanding to a seven-question sequence would greatly improve the ability of capturing this population in the disability population. Table 4 shows with respect to one objective measure – the share of the population captured by diagnosis category – that there is very little difference between the distribution of such conditions in our work-activity-only population (C) and either the 6QS or the 6QS-only population. This is evidence that including the work-activity question in the 6QS would not only increase the share of those receiving Social Security benefits based on their disability but would do so without biasing the sample in this regard.

The most difficult diagnoses to objectively measure are the two largest diagnoses categories in Table 4 – "Mental Disorders other than intellectual disabilities" and "Musculoskeletal System."[5] Table 4 shows that the share of beneficiaries captured in the self-reported disability populations not only do not vary with respect to where they fall in rank order, but are quite close in the percentage of the population they capture in each diagnosis category. There is no evidence that those in the work-activity-only population even in these most difficult to diagnose categories are much different from those in the 6QS or 6QS-only populations.

---

[5]The Social Security Advisory Board [15] found that Social Security determination policy changes in the 1980s stemming from court cases and legislation directly affected how decision makers determined eligibility for applicants who claimed to have musculoskeletal and mental impairments and since then vocational evaluations (which are only required when applicants are not granted eligibility based solely on their medical condition) are more likely to be required for cases where musculoskeletal and mental impairments are alleged.

4.2. How each of the seven CPS questions capture the Social Security disability benefit population

We take further advantage of the added information provided by the SSA administrative records in the 2009 matched-CPS dataset in this section. In Tables 3a, 3b, and 4 we disaggregated our sample of people receiving Social Security benefits based on their disability by the nine largest diagnoses groups and one "other" group containing the rest, and placed them in rows from highest to lowest population in the diagnosis group. We do so again in Table 5 but also include sub-diagnosis categories.

We are now interested in how each of the seven questions in the CPS data are able to capture these people overall and by diagnosis group. So, in this case we look not only at the six questions collectively in the 6QS category (A + B) but also individually. We compare each of the six questions' ability to capture our Social Security beneficiary population and how well they do compared to our work-activity question (B + C) in column 10.

Row 1 of Table 5 reports each of the CPS question's capture value for the entire 9.5 million people represented in our CPS-linked administrative sample. The value in the first column of the first row simply repeats the value for the 6QS in the first row of Table 3b – the 6QS only captures 66.3% of the population. The next six columns show the ability of each of the four function-based questions and the two activity-based questions to capture this population. The next column reports the combined capture rate of these two activity-based questions. The next column reports the capture rate of the work-activity question and the final column reports the capture rate of the seven questions.

The first two function-based questions are related to hearing (7.1%) and vision (8.4%). They are the least able to capture Social Security disability beneficiaries over all. But this is primarily because deafness and blindness are relatively rare in this population. In contrast, when you look at the diagnosis row subcategory "Deafness," the hearing question captures 83.8% of this population. Likewise, the vision question captures 75.8% of the population in the diagnosis row subcategory "Visual Disturbances" and 70.9% in the "Blindness and Low Vision" subcategory. All these subcategory capture rates are substantially higher than the overall capture rate of 66.3% for all 6QS values.

A more subtle impact of these two questions can be seen by looking at the 6QS column (A + B) and comparing its capture rate with that of the work limitation question (B + C). As we saw in Table 3b, in all cases except the "Nervous System and Sense Organs" diagnosis category, the work-limitation question captured a greater share of our Social Security population.

This pattern is repeated for the nine general categories in Table 5 and it is in part because of the ability of the hearing and vision questions to disproportionately capture their primary diagnosis subcategories. But in four general diagnosis categories ("Mental disorders other than intellectual" and "Nervous system and sense organs,"

Table 5

Percentage of Social Security program beneficiaries[a] captured, by diagnosis group and type of disability population measure using the 2009 Matched SSA/CPS-ASEC[b]

Diagnosis group/select diagnosis (diagnosis code(s))	Six question Sequence	Hearing	Vision	Cognitive	Ambulatory	Self-care	Ind. Living	Self-care and/or Ind. Living.	Work activity	Seven question sequence
All groups	66.3	7.1	8.4	31.0	43.0	15.8	31.5	34.6	78.0	89.3
Mental diagnoses other than mental retardation (290–315.2, 315.3–316)	65.9	6.9	6.1	48.4	29.6	11.4	31.2	33.2	78.2	89.8
Other organic psychotic conditions (chronic) (all 294)	72.2	11.1	7.7	56.3	35.2	11.2	28.4	30.3	71.1	93.5
Schizophrenic disorders (all 295)	64.9	4.0	6.3	55.5	16.8	12.1	31.0	32.7	76.7	88.7
Episodic mood disorders (all 296)	64.4	6.5	5.7	44.3	31.8	10.4	29.0	31.3	81.4	89.4
Neurotic disorders (all 300)	64.7	5.2	1.8	45.7	29.9	8.5	37.7	37.9	79.6	91.8
Personality disorders (all 301)	73.4	16.5	15.8	41.6	48.7	33.2	43.2	51.8	60.3	82.7
Physiological malfunction arising from mental factors (306)	77.9	29.3	15.9	46.8	63.8	15.9	48.3	48.3	47.6	77.9
Diseases of the musculoskeletal system and connective tissue (710–739)	66.1	7.2	4.0	17.2	58.9	16.4	23.5	28.1	78.6	89.4
Diffuse diseases of connective tissue (710)	76.3	8.7	5.1	41.5	65.4	7.9	23.4	30.2	82.8	95.4
Rheumatoid arthritis and other inflammatory polyarthropathies (714)	75.1	2.2	4.1	15.1	69.8	38.0	31.0	48.0	83.4	93.1
Osteoarthritis and allied disorders (715)	74.1	8.6	5.5	20.1	67.7	17.7	30.6	32.9	81.1	92.9
Other and unspecified arthropathies (716)	58.5	5.0	3.3	13.4	52.1	10.2	8.8	12.7	62.0	68.8
Other and unspecified disorders of back (all 724)	64.7	7.7	4.1	16.3	57.1	14.7	22.6	26.3	77.4	88.4

...

...

...

...

Table 5, continued

Diagnosis group/select diagnosis (diagnosis code(s))	Disability population measure									
	Six question Sequence	Hearing	Vision	Cognitive	Ambulatory	Self-care	Ind. Living	Self-care and/or Ind. Living.	Work activity	Seven question sequence
Disorders of muscle, ligament, and fascia (728)	50.3	5.7	1.2	16.0	45.1	11.7	10.1	15.7	84.3	89.4
Other disorders of bone and cartilage (733)	50.6	3.3	0.0	22.1	45.0	37.1	34.1	44.9	95.6	100.0
Diseases of the nervous system and sense organs (320–389)	82.7	12.5	23.9	22.7	54.4	28.6	43.7	48.2	80.2	93.5
Other cerebral degenerations (331)	89.7	0.0	13.0	54.0	81.8	38.6	67.4	68.7	71.6	100.0
Other diseases of spinal cord (incl. syringomyelia and syringobulbia) (336)	84.1	9.9	0.0	12.8	84.1	69.4	61.8	70.9	80.0	84.1
Multiple sclerosis (340)	89.0	0.3	17.2	37.1	74.9	41.7	59.3	61.5	85.9	96.4
Infantile cerebral palsy (all 343)	85.8	4.6	11.0	30.2	74.3	43.1	53.3	56.2	71.4	95.8
Epilepsy (345)	64.4	6.2	4.2	37.3	25.2	24.4	38.6	41.2	88.7	93.9
Migraine (346)	51.6	3.6	22.0	36.5	33.3	0.0	18.2	18.2	94.6	94.6
Other and unspecified disorders of the nervous system (349)	74.9	1.2	4.5	24.6	66.4	28.4	41.6	48.2	91.3	91.9
Inflammatory and toxic neuropathy (all 357)	92.6	7.2	21.6	11.0	79.6	34.7	54.5	67.7	87.7	96.7
Muscular dystrophies and other myopathies (359)	94.8	9.0	4.8	39.3	73.5	53.3	28.9	53.3	90.5	100.0
Visual disturbances (368)	81.2	9.9	75.8	0.0	5.3	0.0	50.7	50.7	74.6	93.6
Blindness and low vision (369)	78.9	7.6	70.9	12.3	24.7	12.1	35.4	35.4	82.3	91.8
Deafness (all 389)	87.0	83.8	9.3	6.8	8.4	0.0	8.7	8.7	66.0	87.0
Intellectual disabilities (317–319.5, 319.6–320)	64.5	6.5	6.0	49.8	19.1	13.3	41.0	41.2	76.9	88.8
Mild intellectual disability (317)	47.1	0.0	0.0	37.8	9.3	1.3	19.6	19.6	66.4	74.4

Table 5, continued

Diagnosis group/select diagnosis (diagnosis code(s))	Disability population measure									
	Six question Sequence	Hearing	Vision	Cognitive	Ambulatory	Self-care	Ind. Living	Self-care and/or Ind. Living	Work activity	Seven question sequence
Other specified intellectual disability (all 318)	65.0	6.4	6.2	50.3	19.4	13.6	41.7	41.9	77.0	89.0
Unspecified intellectual disability (319.5)	78.9	1.7	5.6	62.8	33.7	2.3	37.5	37.5	73.2	92.7
Circulatory system (390–459)	64.6	5.1	10.1	21.5	51.6	15.6	31.2	33.3	77.2	88.5
Essential hypertension (401)	80.5	22.9	16.8	27.8	65.1	20.8	42.9	42.9	65.0	90.6
Other forms of chronic ischemic heart disease (414)	61.1	2.8	7.4	19.9	43.7	10.2	25.4	27.1	70.7	87.8
Cardiomyopathy (425)	55.7	1.9	18.1	29.6	45.6	9.9	29.7	36.2	92.9	96.1
Heart failure (428)	38.0	0.0	3.1	10.3	34.4	11.5	28.7	30.4	82.3	89.0
Late effects of cerebrovascular disease (438)	72.5	8.3	7.1	29.6	60.2	27.9	41.7	44.4	77.7	87.9
Other peripheral vascular disease (443)	82.3	0.0	2.0	14.5	75.1	21.1	36.6	40.7	96.4	96.8
Hemorrhoids (454)	45.7	0.0	0.0	0.0	45.7	0.0	25.1	25.1	71.1	71.1
Other disorders of circulatory system (459)	53.7	1.9	10.1	0.0	53.7	7.9	20.6	20.6	56.5	76.2
Injury and poisoning (800–999)	66.6	3.2	7.6	18.0	57.2	32.6	31.5	42.1	82.3	92.6
Fracture of vertebral column with spinal cord injury (806)	76.2	0.0	0.0	8.7	76.2	55.6	48.7	58.6	96.9	97.5
Ill-defined fractures of upper limb (818)	36.1	3.2	12.1	26.5	32.0	16.3	6.9	23.2	74.2	77.2
Other, multiple, and ill-defined fractures of lower limb (827)	66.6	3.4	12.6	20.2	55.4	25.7	19.7	34.6	88.9	97.1
Intracranial injury of other and unspecified nature (854)	56.8	8.5	17.3	26.2	32.3	26.8	24.0	26.8	79.1	89.5
Late effects of musculoskeletal and connective tissue injuries (905)	65.3	9.5	0.0	18.8	48.7	21.2	27.3	31.0	78.9	94.7

Table 5, continued

Diagnosis group/select diagnosis (diagnosis code(s))	Disability population measure									
	Six question Sequence	Hearing	Vision	Cognitive	Ambulatory	Self-care	Ind. Living	Self-care and/or Ind. Living	Work activity	Seven question sequence
Late effects of injuries to the nervous system (907)	66.4	0.0	12.6	23.6	61.9	34.9	52.7	57.8	80.3	88.6
Endocrine, nutritional/metabolic diseases, and immunity disorders (240–248, 249–279)	66.6	6.1	13.0	20.1	60.3	16.2	34.3	37.3	83.0	91.1
Other (248)	68.9	12.0	4.4	23.6	55.1	21.0	36.7	39.4	78.2	90.0
Diabetes mellitus (all 250)	68.6	7.6	19.4	20.8	62.2	13.3	38.3	39.0	80.8	89.1
Obesity and other hyperalimentation (278)	68.0	5.9	8.1	21.1	61.9	22.2	35.1	40.9	83.5	93.1
Respiratory system (460–519)	57.8	5.7	6.7	14.1	48.6	13.1	29.4	31.9	83.7	93.3
Emphysema (492)	83.6	1.2	0.0	7.0	76.4	25.1	45.0	45.0	100.0	100.0
Asthma (493)	52.5	2.8	9.9	17.7	49.5	13.7	24.5	24.5	72.2	84.7
Chronic airway obstruction, NOS (incl. COPD, NOS) (496)	50.3	8.9	7.1	14.1	37.6	8.6	23.2	27.7	87.5	95.6
Neoplasms (140–239)	49.1	2.2	4.6	16.4	39.5	9.8	22.9	24.5	70.9	78.7
Malignant neoplasm of trachea, bronchus, and lung (162)	53.4	0.0	0.0	14.1	39.3	0.0	1.1	1.1	74.7	86.7
Malignant neoplasm of female breast (174)	46.8	0.0	5.0	12.5	27.9	2.7	20.0	22.7	83.2	88.2

[a] SSDI status is based on whether an individual was (a) in the MBR file, (b) receiving benefits on the basis of disability (as opposed to retirement or survivorship), and (c) had a current payment. SSI status is based on whether an individual was (a) in the SSR file, (b) under age 65, and (c) has a current payment.

[b] The 2009 CPS-ASEC file was matched to the SSA Numident file. Numident records were matched for 90,001 of the 102,726 non-institutionalized persons ages 25–61 in the 2009 CPS-ASEC sample – a match rate of 87.6%. Hence 12,725 or 12.5% of persons ages 25–61 in the 2009 CPS-ASEC sample were not matched to Numident records. The CPS-ASEC sample weight was rescaled based on gender and age to account for non-matching sample and reestablish population size estimates.

"Intellectual disabilities," and "Circulatory system") there are sub-diagnosis categories where the 6QS outperforms the work-limitation question in capturing our Social Security beneficiary population. In virtually all those cases the "Hearing" and/or "Vision" questions disproportionately capture Social Security beneficiaries.

The third function-based question is related to cognitive difficulties. It captures 31.0% of our overall Social Security population. But this number is somewhat less impressive than it seems since it is intended to capture the largest and third largest general diagnosis categories. While it does not do as well at capturing "Mental Disorders other than intellectual disabilities" (48.4%) and "Intellectual Disabilities" (49.8%) as "Hearing" and "Vision" do in their primary subcategories, it does disproportionately capture this targeted group.

The fourth function-based question is related to physical mobility (walking and climbing stairs) and it captures the largest (43.0%) part of our overall Social Security disability population. Unlike the other three function-based questions, its capture rate is much more diffuse over all physical diagnoses categories with highs of 60.3% for the "Endocrine et al." category and 58.9% for the "Musculoskeletal et al." category and its capture rates are in the 40- and 50-percent range for the rest. In contrast, its capture rates are much lower for the mental categories – "Mental Disorders other than intellectual disabilities" (29.6%) and "Intellectual Disabilities" (19.1%).

Hence, each of these four function-based questions have capture rates that are higher for our Social Security populations whose diagnosis categories are most closely associated with the function they are trying to identify. This is evidence that each of these four function-based questions has some success in identifying their intended populations with disabilities. Yet collectively they only occasionally exceed the capture rate of the work-activity question.

This last point is important evidence that including a work-activity question in addition to the four function-based questions in the 6QS would substantially increase captures in the Social Security population. This is much less the case with respect to the two activity-related questions in the 6QS, either individually or collectively. They focus on activities of daily living (dressing or bathing) and instrumental activities of daily living (doing errands alone such as visiting a doctor's office or shopping) that require both physical and cognitive skills. But as can be seen across almost all sub-diagnosis categories, the physical mobility-function-based question captures a greater percentage of our Social Security population than either or both these activity-based questions in the diagnosis categories that are physical. Likewise, the cognitive-function-based question captures a greater percentage of our Social Security disability population than either or both of these activity-based questions in the diagnosis categories that are mental. In two of the rare sub-diagnosis cases where this does not occur, "Visual Disturbances" and "Deafness," the vision and hearing-function-based questions surpass them. It is never the case that these two activity-based questions individually or collectively outperform the work-activity question.

Table 5 suggests that for the one group that obviously should be included in any "true population" with disabilities, the four function-based questions in the 6QS do

a reasonable job of capturing the part of our Social Security population whose diagnosis is most closely related to the population targeted by the question. It is far less clear how well the two activity-based questions do in this regard. Unlike the work-activity question that systematically increases the capture rate across almost all diagnosis subpopulations, these two questions almost never surpass the capture rate of the appropriate function-based question for these subpopulations.

5. Reweighting disability samples and possible effects

Our findings suggest that Keller-McNulty was right, and that a work-activity question should be added to the 6QS to better capture SSDI (Disabled Workers, Disabled Window(ers) and DAC) and SSI-Disabled Adult beneficiaries. But until this happens, at a minimum, the Census Bureau and the Department of Labor who are responsible for the development of the CPS, as well as individual researchers using these data, should consider reweighting their disability samples based on the 6QS, the work-activity question, and the union of these two ways of capturing the disability population to better capture the share of those question-based populations that are receiving SSDI and SSI benefits. They should also do so to better estimate the economic outcomes of their overall disability population.

In Table 6 we show how sensitive labor force participation, employment, and poverty rates are to two such alternative reweighting methods. Because most researchers will not have easy access to the administrative records data, in this example, we identify Social Security beneficiaries using only public use self-reported information on SSDI and SSI benefit status in the 2009 Public Release CPS-ASEC.

The first panel of Table 6 reports population rates for these three economic outcome variables for the 6QS-based population, the work activity-based population, and for the union of the two populations. In the next panel we show rates when we reweight beneficiaries observed in the population to adjust for those who reported receiving SSDI or SSI-Disabled Adult benefits but who are not included in the population with disabilities. Doing so assumes that these missing observations have on average the same labor force participation, employment, and poverty rates as those we observe. In the last panel, we simply add to the disability population all missing observations who reported receiving SSDI or SSI-Disabled Adult benefits.

As a result, in all our disability populations, labor force participation and employment rates fall and poverty rates rise. But the greatest change is in the 6QS-based population since it misses the greatest share of the SSDI and SSI population.

6. Discussion

The Affordable Care Act of 2010 mandates the establishment of standards for the collection and dissemination of health statistics by disability status. In response, the

Table 6

Economic outcomes of non-institutionalized persons ages 25–61 under the assumption that each disability population measure type captured all SSDI and SSI participants using two methods,[a,b] by disability measure, using the 2009 Public Release CPS-ASEC

Economic variable	Disability population measure		
	Six question sequence (A+B)	Work limitation (B+C)	Seven question sequence (A+B+C)
Employment rate	31.1	17.3	29.0
Labor force participation rate	35.9	21.3	34.2
Poverty rate	25.9	29.2	26.2
	Method 1: Reweighting Identified SSDI and SSI participants to reflect total participants		
Employment rate	26.8	16.3	28.3
Labor force participation rate	30.8	20.1	33.4
Poverty rate	26.7	29.4	26.3
	Method 2: Adding in missed SSDI and SSI participants		
Employment rate	27.3	17.0	28.7
Labor force participation rate	31.6	20.8	33.8
Poverty rate	27.0	29.6	26.5

[a]SSDI participation is based on responses to the following question: Did (you/name) receive Social Security? [If yes] what were the reasons (you/name) (was/were) getting Social Security income last year? [retired; disabled (adult or child); widowed; spouse; surviving child; dependent child; on behalf of surviving, dependent, or disabled child(ren); other (adult or child)]. Respondents were allowed two reasons. SSI participation is based on responses to the following question: Did (you/name) receive SSI? [If yes] what were the reasons (you/name) (was/were) getting Supplemental Security Income last year? [disabled (adult or child); blind (adult or child); on behalf of a disabled child; on behalf of a blind child; other (adult or child)]. Respondents were allowed two reasons.
[b]We utilize two approaches: (a) reweight beneficiaries observed in each disability population to reflect the total number of beneficiaries and (b) add in the missing beneficiaries into the disability populations identified by each measure. Both will yield the same population and prevalence rates, however their employment rates will differ because the reweighting approach assumes the employment rate of the observed beneficiaries in the disability population is the same as the employment rate of the missed beneficiaries. The "add in" approach does not need to make such an assumption because the employment of the missed beneficiaries is known and used.

U.S. Department of Health and Human Services recommended that the 6QS, first used in the ACS and more recently in the CPS, be the minimum data collection standard for survey questions on disability status [9].

However, none of these six questions directly relates to work-activity limitations. We examine the consensus view as expressed in HHS [9]. In doing so, we provide evidence that the lack of a work-activity question in the 6QS results in its inability to capture a substantial portion of the population with disabilities relevant to key U.S. disability policies and programs.

Using linked 2009 CPS-ASEC/SSA records data, we find that this 6QS captures only 66.3% of those receiving Social Security benefits based on their disability (a group that presumably should be captured as a subpopulation of any more general disability population). Furthermore, substantial portions of Social Security recipients within diagnostic groups are not captured by the 6QS.

When we add a work-activity question to the 6QS, we increase our Social Security disability population captured by 23.1 percentage points, for a total of 89.3%. Hence a seven-question-sequence would capture a larger share of a subpopulation that should be in any population with disabilities – SSDI and SSI-Disabled Adult beneficiaries.

Importantly, we also provide evidence that this added population is not much different with respect to its distribution of diagnoses than the 6QS or even the 6QS-only disability populations. Hence, including it with the 6QS or even the 6QS-only population does not change the distribution of diagnoses of those captured by the larger disability population – the one additional piece of objective information based on administrative records data we have on the characteristics of this missing piece of the disability populations these questions capture.

This similarity with respect to the distribution of diagnoses captured is in sharp contrast to the major differences that failing to capture SSDI and SSI beneficiaries using these alternative definitions of the disability population with CPS public use data has on estimated labor force participation, employment, and poverty rates. As discussed in Table 6, a short-term way to adjust for these Type 2 errors could be reweighting the disability samples based on how well they capture SSDI/SSI-Disabled Adult recipients. These reweighting exercises provide a first approximation of the degree that current CPS-based statistics based on the 6QS and the work-activity questions overstate the labor force participation and employment rates and understate the poverty rates of working-age people with disabilities by their failure to capture all those receiving SSDI and SSI benefits in their disability definitions.

But we conclude that the absence of a work-activity question in the 6QS makes it a fundamentally flawed data standard for survey questions on disability. The long run solution is to belatedly follow Keller-McNulty's [12] advice and make the technical and methodological adjustments to a work-activity question so that it can be included in any minimum data collection standard for survey questions on disability status.

Acknowledgements

References

[1]   R.V. Burkhauser, M.C. Daly, A.J. Houtenville and N. Nargis, Self-reported work limitation data: What they can and cannot tell us, *Demography* **39**(3) (2002), 541–555.

[2]   R.V. Burkhauser, A.J. Houtenville and J.R. Tennant, Capturing the elusive working-age population with disabilities: Reconciling conflicting social success estimates from the Current Population Survey and the American Community Survey, *Journal of Disability Policy Studies* **24**(4) (2014), 195–205.

[3]   Current Population Survey, *Design and Methodology, Technical Paper 66* (2006), http://www.census.gov/prod/2006pubs/tp-66.pdf.

[4]   Current Population Survey, *Annual Social and Economic (ASEC) Supplement Technical Documentation* (2009), http://www.census.gov/apsd/techdoc/cps/cpsmar09.pdf.

[5]   Current Population Survey, *Frequently Asked Questions About Disability Data* (2009), http://www.bls.gov/cps/cpsdisability_faq.htm.

[6]   J.L. Czajka, J. Mabli and S. Cody, *Sample Loss and Survey Bias in Estimates of Social Security Beneficiaries: A Tale of Two Surveys*. Washington, DC: Mathematica Policy Research, 2008, http://www.mathematica-mpr.com/publications/pdfs/samplelosssurvey.pdf.

[7]   P.S. Davies and T.L. Fisher, Measurement issues associated with using survey data matched with administrative data from the Social Security Administration, *Social Security Bulletin* **69**(2) (2009), 1–12.

[8]   V.A. Freedman, E. Crimmins, R.F. Schoeni, B. Spillman, H. Aykan, E. Kramarow, K. Land, J. Lubitz, K. Manton, L.G. Martin, L.D. Shinberg and T. Waidmann, Resolving inconsistencies in old-age disability trends: Report from a technical working group, *Demography* **41**(3) (2004), 417–441.

[9]   HHS (U.S. Department of Health and Human Services), *Implementation Guidance on Data Collection Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status* (2011), http://aspe.hhs.gov/datacncl/standards/ACA/4302/index.pdf.

[10]   A.J. Houtenville, D.C. Stapleton, R.R. Weathers II and R.V. Burkhauser, eds., *Counting Working-age People with Disabilities: What Current Data Tell Us and Options for Improvement*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 2009.

[11]   A.M. Jette and E. Badley, Conceptual issues in the measurement of work disability, in: *The Dynamics of Disability: Measuring and Monitoring Disability for Social Security Programs*, G.S. Wunderlich, D.P. Rice and N.L. Amado, eds., Washington DC: National Academy Press, 2002, pp. 183–210.

[12]   S. Keller-McNulty, Reference: American Community Survey and disability. Letter to Dr. Lewis Kincannon, Director, U.S. Census Bureau, reproduced in *AmStat News*, June 2006, p. 4.

[13]   N.A. Mathiowetz, Methodological issues in the measurement of work disability, in: *The Dynamics of Disability: Measuring and Monitoring Disability for Social Security Programs*, G.S. Wunderlich, D.P. Rice and N.L. Amado, eds., Washington DC: National Academy Press, 2002, pp. 211–240.

[14]   K. Miller and T.J. DeMaio, *Report of Cognitive Research on Proposed American Community Survey Disability Questions* (2006), Washington DC: US Census Bureau.

[15]   Social Security Advisory Board, *Aspects of Disability Decision Making: Data and Materials* (2012), http://ssab.gov/Publications/Disability/GPO_Chartbook_FINAL_06122012.pdf.

[16]   D. Stapleton, T. Honeycutt and B. Schechter, Out of sight, out of mind: Including group quarters residents with household residents can change what we know about working-age people with disabilities, *Demography* **49**(1) (2012), 267–289.

[17]   G. Swanson, L. Carrothers and K. Mulhorn, Comparing disability survey questions in five countries: A study using ICF to guide comparisons, *Disability and Rehabilitation: An International Multidisciplinary Journal* **25** (2003), 665–675.

[18]   World Health Organization, *International Classification of Disability, Health and Functioning* (2001), Geneva: World Health Organization.

Appendix Table 1. Disability, employment, and program participation questions in the Current Population Survey

Question (Survey)	Question wording
Disability question	
Hearing difficulty (CPS-BMS)	Is anyone deaf or does anyone have serious difficulty hearing?
Vision difficulty (CPS-BMS)	Is anyone blind or does anyone have serious difficulty seeing even when wearing glasses?
Mental difficulty (CPS-BMS)	Because of a physical, mental, or emotional condition, does anyone have serious difficulty concentrating, remembering or making decisions?
Physical difficulty (CPS-BMS)	Does anyone have serious difficulty walking or climbing stairs?
Self-care difficulty (CPS-BMS)	Does anyone have difficulty dressing or bathing?
Independent living difficulty (CPS-BMS)	Because of a physical, mental or emotional condition, does anyone have difficulty doing errands alone such as visiting a doctor's office or shopping?
Work-activity limitation (CPS-ASEC)	Does anyone in this household have a health problem or disability which prevents them from working or which limits the kind or amount of work they can do?
Employment question	
Currently employed (CPS-BMS)	Last week, did [person] do any work for either pay or profit?
Worked at least 52 hours in the prior calendar year (CPS-ASEC)	Work hours $\geqslant$ 52. To construct this variable, use the following two questions: (1) During [the previous calendar year] in how many weeks did [person] work even for a few hours? Include paid vacation and sick leave as work, and (2) In the weeks that [person] worked [the previous calendar year], how many hours did [person] usually work per week?
Worked full-time, full-year in the prior calendar year (CPS-ASEC)	Work hours per week $\geqslant$ 35 and work weeks per year $\geqslant$ 50. To construct this variable, use the following two questions: (1) During [the previous calendar year] in how many weeks did [person] work even for a few hours? Include paid vacation and sick leave as work, and (2) in the weeks that [person] worked [the previous calendar year], how many hours did [person] usually work per week?
Program participation question	
Social Security Disability Insurance (CPS-ASEC)	Did (you/name) receive Social Security? [If yes] what were the reasons (you/name) (was/were) getting Social Security income last year? [retired; disabled (adult or child); widowed; spouse; surviving child; dependent child; on behalf of surviving, dependent, or disabled child(ren); other (adult or child)]. Respondents were allowed two reasons.
Supplemental Security Income (CPS-ASEC)	Did (you/name) receive SSI? [If yes] what were the reasons (you/name) (was/were) getting Supplemental Security Income last year? [disabled (adult or child); blind (adult or child); on behalf of a disabled child; on behalf of a blind child; other (adult or child)]. Respondents were allowed two reasons.

Appendix Table 1A. ACS Work-Activity Question (Dropped in 2008)

Question (Survey)	Question wording
Work-activity question (ACS)	Because of a physical, mental, or emotional condition lasting six months or more, does this person have any difficulty in doing any of the following activities: working at a job or business?