

RESEARCH ARTICLE

# Disability policies and public views on work disability: A comparative analysis using anchoring vignette data

Na Yin<sup>1,2,3\*</sup> and Frank Heiland<sup>1,2,3</sup>

<sup>1</sup>Marx School of Public and International Affairs, Baruch College, The City University of New York (CUNY), New York, NY, USA

<sup>2</sup>CUNY Institute for Demographic Research, New York, NY, USA

<sup>3</sup>The CUNY Graduate Center, New York, NY, USA

**Abstract:** Using data on disability vignettes from representative surveys in the U.S. and seven European countries, we conduct a comparative analysis of disability policies and public views on work limitations. We hypothesize that program characteristics are related to individuals' perceptions about work limitations. Looking at how respondents across countries characterize identical disability vignettes, we find evidence that disability policy dimensions such as policy coverage, medical assessment, and vocational assessment strongly predict disability perceptions. We illustrate the results in a series of counterfactual policy simulations. Our findings have implications for policy design and delivery. The anchoring vignette approach may also be useful in a wide range of comparative policy studies.

**Keywords:** *comparative analysis of disability policies; disability perception; anchoring vignette approach*

## ARTICLE INFO

Received: September 6, 2016  
Accepted: October 24, 2016  
Published Online: November 3, 2016

### \*CORRESPONDING AUTHOR

Na Yin, Marx School of Public and International Affairs, Baruch College & CUNY Institute for Demographic Research & The CUNY Graduate Center, One Bernard Baruch Way D901, New York, NY 10010, USA;  
Na.Yin@baruch.cuny.edu

### CITATION

Yin N and Heiland F (2017). Disability policies and public views on work disability: A comparative analysis using anchoring vignette data. *International Journal of Population Studies*, 3(1): 42–63.  
doi: [10.18063/IJPS.2017.01.005](https://doi.org/10.18063/IJPS.2017.01.005).

**Copyright:** © 2017 Na Yin and Frank Heiland. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1 Introduction

### 1.1 Motivation and Research Questions

Population aging has led to increases in activity limitations among the pre-retirement people across countries (Martin and Schoeni, 2014; Global Burden of Disease Study 2013 Collaborators, 2015). It has presented significant challenges to societies trying to maintain a productive work force and fund pay-as-you-go programs supporting older individuals such as social security. With recent reforms in many countries reducing the generosity of public pensions (e.g., by increasing the age at which workers are eligible to draw full benefits), older workers have increasingly sought assistance from public disability programs (Duggan, Singleton and Song, 2007; GAO, 2010).

Policymakers seek to both ensure the wellbeing of older workers with health impairments and to improve the efficiency and sustainability of disability systems. For example, many European countries now emphasize the importance of supporting and encouraging work among those with health limitations (OECD, 2010). Similarly, U.S. disability policy, which has traditionally focused on providing income to individuals who cannot do any work, is experimenting with providing work incentives for individuals with health limitations. Recent examples include the Ticket to Work Program and the Benefit Offset National Demonstration Projects (see Livermore *et al.* (2013) for a summary of the main findings from evaluating the Ticket to Work Program, and see Weathers and Hemmeter (2011), for preliminary findings from the Benefit Offset National Demonstration pilot projects). Appendix A provides detailed discussion about the benefit structure of OECD countries' disability programs and their policy lessons.

An active literature is devoted to estimating the impact of disability reform on behaviors such as labor force participation and program application (e.g., Gruber, 2000; Mullen and Staubli, 2016; Wise, 2017). We argue that there is an important pathway through which disability policy may influence individuals that has been largely overlooked: individuals' perceptions regarding what constitutes a work disability. In this paper, we provide a comparative analysis of disability policy in the U.S. and seven European countries and investigate whether cross-country differences in disability policies are linked to different views held by the public on work limitations.

We hypothesize that long-standing institutional differences (e.g., coverage and benefit generosity) between the public disability insurance systems in Europe and the U.S. affect residents' perceptions of work limitations, and in turn shape beliefs regarding disability. For instance, since many European disability programs differentiate by severity (offering partial benefits) while the U.S. only recognizes those with the most severe work limitations as disabled (offering full support), Americans may consider a person with a health condition that limits but does not prevent work as not work-disabled while Europeans may describe him or her as moderately work-disabled. As another example, in countries with better accessibility, it is more common for people to see disabled individuals on a daily basis performing regular activities whereas in other countries the disabled individuals might not be able to leave their houses as often, which might affect public views on disability.<sup>1</sup>

To further conceptualize the relationship between disability institutions and individuals' perception of work disability, we invoke the concept of cultural beliefs. Culture attracts researchers in many fields, but there has not been a universal definition that applies across disciplines (Bachrach, 2013; Bisin and Thierry, 2010; Giuliano, 2007). When we refer to (disability) culture here, we follow literature in economics (e.g., Guiso, Sapienza and Zingales, 2006) and anthropology (e.g., Boyd and Richerson, 2005) that define culture as "decision-making heuristics or rules of thumb that have evolved to serve our need to make decisions in complex and uncertain environments ... manifest[ing] themselves as values, beliefs, or social norms" (Alesina and Giuliano, 2013, p. 5).

Cultural beliefs have been linked to the institutional environment: People internalize social norms that emerge and develop in specific institutional settings (Alesina and Giuliano, 2013). As a result, differences in policy can affect the prevalence of various types of social norms. (For example, Alesina and Glaeser (2004) while examining the bi-directional relationship between institutions and culture, show that the different attitudes between Americans and Europeans toward the poor can be explained by the relative generosity of the respective welfare states.) Adopting the same conceptual framework, we hypothesize that the disability institutional arrangements in a country will play a role in shaping the social norms about work disabilities, as reflected in the way people assess work limitations.

Previous research has documented substantial variations across countries in disability assessments that cannot be explained by demographic and health characteristics (e.g., Angelini, Cavapozzi and Paccagnella, 2012; Bagod'Uva, O'Donnell and van Doorslaer, 2008; Grol-Prokopczyk, Freese and Hauser, 2011; Kapteyn, Smith and van Soest, 2007; 2009; Murray *et al.*, 2003; Sadana *et al.*, 2002). However, to date, systematic analyses of the potential mechanisms underlying those cross-country differences has been lacking. We seek to fill this void by conducting a comparative analysis of disability policy, investigating how the different components of the disability systems in the U.S. and Europe influence individuals' perceptions of what constitutes work limitations.

To study individuals' perceptions, we look at how differently people characterize a given level of work disability across countries (disability reporting style). Measures of self-assessed disability status, which are commonly available in survey data, are insufficient to conduct this type of analysis because they will reflect both the true

1. We thank an anonymous referee for making this point.

(underlying) level of work disability and a country-specific way of reporting. We take advantage of unique data on disability vignettes from comparable U.S. and European surveys. We identify rating styles (reporting heterogeneity) based on how respondents characterize the severity level of the same vignette person with a given degree of work limitation. This is known as the anchoring vignette approach, which has become a popular empirical strategy in health and social science research (King *et al.*, 2004).

We find evidence consistent with an effect of disability policy generosity on perceptions. Different policy dimensions affect disability classifications in different ways. The most influential policy dimensions are policy coverage, medical assessment, and vocational assessment. Our findings have important implications for disability policy design and delivery.

## 1.2 Disability Policy and Disability Perception

In this section, we compare the disability policies in the United States and seven European countries, and briefly discuss the mechanisms of how individuals perceive work limitations under different policy environments.

We conceptualize public policy as the outcome of a process that balances individuals' wants for service provision with the available societal resources' given preferences (ranking over priorities). In turn, differences in disability policies across countries with similar standards of living should reflect differences in the public's (consensus) views regarding the relative importance and the appropriate type of work disability support.

For simplicity, economists often consider preferences underlying a policy choice as invariant. However, individuals' attitudes regarding the objects of a policy are likely to be influenced by social norms, culture, and tradition as discussed in the Introduction. If so, then we expect policies to also shape individuals' perceptions, especially those that have been in place for extended time periods.

To investigate this hypothesis, we test empirically whether disability policies impact how individuals characterize disability cases across countries. Disability policy may provide knowledge—in the form of references or thresholds—on how to characterize or classify work limitations. For example, in a disability welfare state that traditionally recognizes various extents of work limitations, the society members may rate a health problem on a wider (more differentiated) severity spectrum.

In contrast, in more laissez-faire society disability policy may be very strict, offering little support to work-limited individuals and only to the most severe disability cases that have little residual work capacity. Under such a disability regime, people will get the impression that disability support is very limited and that mild or moderate health impairments are neither defined as “disabling” nor qualifying for public support. As a result, members in these societies likely apply a higher threshold when classifying disability.

OECD (2003) provides a classification for disability compensation policy for all its member countries. For every country, 10 disability policy dimensions are evaluated: coverage, minimum disability level, disability levels for full benefits, maximum benefit level, permanence of benefits, medical assessment, vocational assessment, sickness benefit level, sickness benefit duration, and unemployment benefit level and duration. OECD scores the generosity of each dimension of disability policy on a scale from 0 to 5, with 5 being the highest score (most generous) and 0 (least generous) the lowest. The OECD measures are designed to capture not only the formal disability program rules but also their implementation and administration. [Table 1](#) provides a detailed overview of the classifications of the disability compensation policy dimensions. We also list the eight countries in our analytical sample according to their scoring in each policy dimension. Below, we provide a comparative analysis for each policy aspect.

### 1.2.1 Policy 1: Coverage

Nordic countries generally provide full population coverage for public disability benefit programs, regardless of the individual's work history and household income. In terms of the OECD scoring, Sweden scores 5 for having the most extensive coverage which

**Table 1.** OECD classification for disability compensation policy dimension (based on OECD 2003, Table A2.1)

	<b>5 points</b>	<b>4 points</b>	<b>3 points</b>	<b>2 points</b>	<b>1 point</b>	<b>0 point</b>
<b>Policy 1: Coverage</b>	total population (residents) <i>Sweden</i>	some of those out of the labor force <i>Netherlands</i>	labor force plus means-tested non-contribution scheme <i>Belgium, France, Italy, Spain, U.S.</i>	labor force with voluntary self-insurance <i>Germany</i>	labor force	employees
<b>Policy 2: Minimum Disability Level</b>	0%–25% <i>Germany, Netherlands, Sweden</i>	26%–40% <i>Spain</i>	41%–55%	56%–70% <i>Belgium, France, Italy</i>	71%–85% <i>U.S.</i>	86%–100%
<b>Policy 3: Disability Levels for Full Disability</b>	<50%	50%–61%	62%–73% <i>Belgium, Germany</i>	74%–85% <i>Netherlands, U.S.</i>	86%–99% <i>France, Spain, Sweden</i>	100% <i>Italy</i>
<b>Policy 4: Maximum Benefit Level</b>	$rr \geq 75\%$ , reasonable minimum <i>Netherlands, Sweden</i>	$rr \geq 75\%$ , minimum not specified <i>Spain</i>	$75 > rr \geq 50\%$ , reasonable minimum <i>France, Italy, U.S.</i>	$75 > rr \geq 50\%$ , minimum not specified <i>Germany</i>	$rr < 50\%$ , reasonable minimum <i>Belgium</i>	$rr < 50\%$ , minimum not specified
<b>Policy 5: Permanence of Benefits</b>	strictly permanent <i>Spain</i>	<i>de facto</i> permanent <i>Belgium, U.S.</i>	self-reported review only <i>Netherlands, Sweden</i>	regulated review procedure	strictly temporary, unless fully (= 100%) disabled <i>France, Germany, Italy</i>	strictly temporary in all cases
<b>Policy 6: Medical Assessment</b>	treating doctor exclusively	treating doctor predominantly <i>U.S.</i>	insurance doctor predominantly <i>Germany, Sweden</i>	insurance doctor exclusively <i>Belgium, France</i>	team of experts in the insurance <i>Italy, Netherlands</i>	insurance team and two-step procedure <i>Spain</i>
<b>Policy 7: Vocational Assessment</b>	strict own or usual occupation assessment	reference is made to one's previous earnings <i>Belgium, France, Germany (3.5)</i>	own-occupation assessment for partial benefits <i>Italy, Spain</i>	current labor market conditions are taken into account	all jobs available taken into account, leniently applied <i>Netherlands, Sweden, U.S.</i>	all jobs available taken into account, strictly applied
<b>Policy 8: Sickness Benefit Level</b>	$rr = 100\%$ also for long-term sickness absence	$rr = 100\%$ (short-term); $rr \geq 75\%$ (long-term) sickness absence <i>Germany, Sweden</i>	$rr = 75\%$ (short-term); $rr \geq 50\%$ (long-term) sickness absence <i>Belgium, Italy, Netherlands</i>	$75 > rr \geq 50\%$ for any type of sickness absence <i>France, Spain, U.S.</i>	$rr \geq 50\%$ (short-term); $rr < 50\%$ (long-term) sickness absence	$rr < 50\%$ also for short-term sickness absence
<b>Policy 9: Sickness Benefit Duration</b>	one year or more, short or no wage payment period <i>France</i>	one year or more, significant wage payment period <i>Germany, Spain, Sweden</i>	6–12 months, short or no wage payment period <i>Italy, Netherlands</i>	6–12 months, significant wage payment period <i>Belgium</i>	<6 months, short or no wage payment period	<6 months, significant wage payment period <i>U.S.</i>

	5 points	4 points	3 points	2 points	1 point	0 point
<b>Policy 10: Unemployment Benefit (UE) Level and Duration</b>	DI>UE level, short duration of unemployment	DI>UE level, long duration of unemployment	similar levels, short duration of unemployment	similar levels, long duration of unemployment	DI<UE level, short duration of unemployment	DI<UE level, long duration of unemployment
		<i>Spain</i>	<i>Italy, Sweden</i>	<i>Belgium, France, Germany, Netherlands</i>	<i>U.S.</i>	

Note: rr=replacement rate; DI=Disability benefit

covers its whole population. Many disability systems (e.g., Belgium, France, Italy, Spain, and the U.S.) cover labor force plus a means-tested non-contribution scheme. In the disability systems that cover mainly labor force, five years of work are typically required to establish entitlements (Belgium and France require shorter contribution histories), and some of those countries (e.g., the U.S.) require several or all of the five years of work to have occurred recently. Usually the contribution requirements for sickness benefits are much weaker.

### 1.2.2 Policy 2: Minimum Disability Level

The U.S. federal disability system intends to award benefits only to the individuals who are fully and permanently disabled. It does not allow partial or temporary disabilities. The U.S. scores the lowest by the OECD classification for its strictest “minimum disability level”. To receive any disability benefits, an American must have at least 71%–85% disability level, compared to only 0%–25% for a German, a Dutch, or a Swede; 26%–40% for a Spaniard; and 56%–70% for a Belgian, French or an Italian (we refer interested readers to Yin (2015) for a detailed analysis of the incentives provided by the U.S. all-or-nothing disability system in contrast with the European partial disability system and for a simulation of the effects of introducing partial benefits into the U.S. system on application behavior and employment).

Most countries with such systems, including some Nordic (e.g., Sweden), western (e.g., Germany, and the Netherlands), and central and eastern European countries, offer a full benefit to those assessed to be incapable of work, as well as various partial benefits consistent with reduced work capacity. They offer one (e.g., Germany) to four different levels (e.g., Sweden) of partial benefits and in some cases offer finer gradations (e.g., the Netherlands).

Other countries, including France and Spain, have a quasi-partial benefit for people who are unable to work in their usual occupation and a full benefit only for those unable to work in any occupation. Hence, the capacity threshold is the same for both benefits but the reference is different. The partial benefit is allowed to be supplemented to some extent by earnings from a job in another occupation.

Even European countries that do not offer partial benefits as part of their main disability program, such as Belgium and Italy, have universal sickness programs as a precursor to their long-term disability programs and also more alternative public programs to complement their disability program. For example, while the Italian disability pension scheme awards full benefits only to people totally unable to work, a means-tested disability allowance compensates for partial work capacity loss. We refer interested readers to OECD (2010) for a detailed discussion about OECD countries’ recent experience and policy lessons in reforming their disability programs.

### 1.2.3 Policy 3: Disability Level for Full Disability

Despite the relatively lenient standard used in Europe, compared to the U.S., in awarding disability benefits to people with less severe work limitations (reflected in the lower minimum disability level, higher and more durable sickness benefits, and more generous disability benefits than unemployment benefits), European countries require a higher or similar disability severity level to qualify for full disability benefits.

For example, Sweden requires as high as 86%–99% disability level for full disability benefits while its minimum disability level is only 0%–25%. The U.S. system does not offer benefits for mildly or moderately work-limited individuals. It allows disability benefits only to those with a disability level of 71%–85% or higher.

#### **1.2.4 Policies 4 and 5: Maximum Benefits and Permanence of Benefits**

According to the proportion of the work earnings that can be replaced by the maximum benefit level, Sweden and Netherlands rank the highest with the most generous replacement rate ( $\geq 75\%$ ), and the U.S. sets a moderate replacement rate of 50%–75%, the same as the rate in France and Italy.

The outflow from disability rolls back to employment is almost zero in all the countries, not only in the country where the disability compensation is long-term (e.g., strictly permanent in Spain; *de facto* permanent in Belgium and U.S.) but also in countries where disability benefits are supposed to be temporary (unless for fully disabled) such as in Sweden, France, Germany and Italy. In most countries, periodic legal reviews are required on disability beneficiaries but in practice they rarely occur.

#### **1.2.5 Policies 6 and 7: Medical Assessment and Vocational Assessment**

In most countries, the medical assessment is performed by insurance doctors. However, countries differ in accounting for the opinion of treating doctors. In most of the countries we study, such opinion is not taken into account in disability assessments. Germany and Sweden weigh in the medical evaluation. In the U.S., the disability determination relies on the applicants' treating doctors' opinion predominantly. On this policy measure, the U.S. takes the lead in leniency. While the treating doctors may be better informed about the applicant's medical problems, there are also concerns that the treating doctors are too familiar with the applicant to make an objective assessment.

In terms of vocational assessment, Sweden, like the U.S., takes into account all the jobs available in the national economy, whereas some other countries, such as Belgium and France, use one's previous earnings/occupations as reference in evaluating one's residual earning/work capacity.

#### **1.2.6 Policies 8 and 9: Sickness Benefits**

Sickness benefits in many countries usually target individuals with less severe work limitations and serve as a precursor to long-term disability programs. In the Netherlands, nearly all employees receive 100 percent wage replacement (70 percent sickness benefit topped up by collective bargaining to 100 percent of the wage) during the entire sickness period. Germany and Sweden have a higher benefit rate for sickness than for disability, while in the southern European countries sickness benefits are usually much lower than disability benefits. The U.S. does not have any sickness benefit program although some employers and five states provide short-term disability benefits.

#### **1.2.7 Policy 10: Unemployment Benefits**

A worker who suffers health impairment and job separation but still has residual work capacity may choose to apply for unemployment benefits or partial disability benefits. This is mostly likely the case in European countries whose disability systems offer partial benefits. In the U.S., the disability system awards benefits only to the fully disabled individuals who are not supposed to qualify for any unemployment benefits. In Spain, disability benefits are more generous than unemployment benefits. Sweden offers similar levels of disability benefits and unemployment benefits for a resident but the duration is longer for disability benefits. The U.S. is one of the few developed countries where disability benefit levels are significantly lower than unemployment benefit levels, although the duration for unemployment benefits is rather short.

By summing up the scores in the 10 dimensions discussed above, we obtain an overall policy generosity value for each country. The U.S. ranks as the least generous

disability system (index=21) and Sweden ranks the most generous (index=34). However, countries rank differently in each specific policy dimension. Some policy dimensions, such as Policies 2, 8, 9 and 10, are highly correlated, as evidenced by the magnitudes of the correlation coefficients (0.7 or more between any two). Moreover, these four policy dimensions affect the vignettes' ratings in a similar fashion, as indicated by the correlation coefficients between each policy dimension and the vignettes classifying ([Appendix Table A1](#)). In our estimation and policy simulation shown later, we group these highly correlated policy dimensions to reduce collinearity. More importantly, these four aspects likely reflect how a country treats milder disability cases, as the policies regarding the minimum disability level (Policy 2), sickness benefits (Policies 8 and 9), and unemployment benefits (Policy 10) target partial or temporary disability cases.

The remainder of the paper is structured as follows: In Section 2, we describe the data and the anchoring vignette approach. In Section 3, we present estimation results and policy simulation results. Section 4 provides further discussion about the results and concludes.

## 2 Data and Method

### 2.1 Study Sample

To study individuals' perceptions, we look at how differently people characterize a given level of work disability across countries. Measures of self-assessed disability status, which are commonly available in survey data, are insufficient to conduct this type of analysis because they will reflect both the true level of work disability and reporting styles. We take advantage of unique data on disability vignettes from comparable U.S. and European surveys, and use vignette data to study reporting heterogeneity. A vignette describes the work limitation of a hypothetical person. A respondent is asked to evaluate the severity of the vignette work limitation on the same five-point scale used for their own health assessment. Since the vignettes are identical for all the respondents, the differences in respondents' evaluations must be due to different reporting styles. We hypothesize that the scale that the respondents use to classify the severity of a given vignette character's work limitation is a function of the country's disability policy, particularly a severity classification scale used by their country's disability system.

We use the 2004 wave of the Health and Retirement Study (HRS), a bi-annual panel with a representative sample of the U.S. population aged over 50 and their spouses. It has been conducted by the University of Michigan since 1992. The information collected includes health, socio-economic status, and social program participation. We use a subsample of respondents who first completed a face-to-face interview and later completed a leave-behind questionnaire that consists of a series of work disability vignettes. We use the 2004 wave because it is the only year in the panel when a vignette questionnaire was given to a random subgroup of respondents.

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a longitudinal dataset on European citizens of aged 50 and older and their spouses. We use the 2004 wave of the survey. SHARE was purposely modeled after the HRS and follows a common set-up across all countries with the goal of facilitating cross-country research. For a subset of countries that agreed to participate, SHARE included a set of self-assessments and vignette questions on work limitations as part of a drop-off questionnaire. The eight countries that participated in this vignette experiment were Belgium, France, Germany, Greece, Italy, the Netherlands, Spain, and Sweden. The work disability vignettes were identical to the work disability vignettes in the HRS leave-behind questionnaire. In our analysis, we exclude Greece because the comparable index for the disability policy generosity is not available in the OECD report.

## 2.2 Measurement

The work disability vignettes describe work limitation of a hypothetical character in three domains: pain, depression, and cardiovascular health. In each domain, several vignette questions are asked. We use the nine vignettes common to the HRS and SHARE surveys. The text for all the vignettes is provided in [Table 2](#).

**Table 2.** Vignettes text in the HRS/SHARE questionnaires

---

<b>Pain Vignettes:</b>	
1	[Name] has almost constant pain in her back and this sometimes prevents her from doing her work.
2	[Name] suffers from back pain that causes stiffness in her back especially at work but is relieved with low doses of medication. She does not have any pains other than this generalized discomfort.
3	[Name] has pain in his back and legs, and the pain is present almost all the time. It gets worse while he is working. Although medication helps, he feels uncomfortable when moving around, holding and lifting things at work.
<b>Cardiovascular Vignettes:</b>	
1	[Name] has had heart problems in the past and he has been told to watch his cholesterol level. Sometimes if he feels stressed at work, he feels pain in his chest and occasionally in his arms.
2	[Name] has been diagnosed with high blood pressure. His blood pressure goes up quickly if he feels under stress. Tom does not exercise much and is overweight.
3	[Name] has undergone triple bypass heart surgery. He is a heavy smoker and still experiences severe chest pain sometimes.
<b>Depression Vignettes:</b>	
1	[Name] feels worried all the time. She gets depressed once a week at work for a couple of days in a row, thinking about what could go wrong and that her boss will disapprove of her condition. But she is able to come out of this mood if she concentrates on something else.
2	[Name] has mood swings on the job. When she gets depressed, everything she does at work is an effort for her and she no longer enjoys her usual activities at work. These mood swings are not predictable and occur two or three times during a month.
3	[Name] generally enjoys her work. She gets depressed every three weeks for a day or two and loses interest in what she usually enjoys but is able to carry on with her day-to-day activities on the job.

---

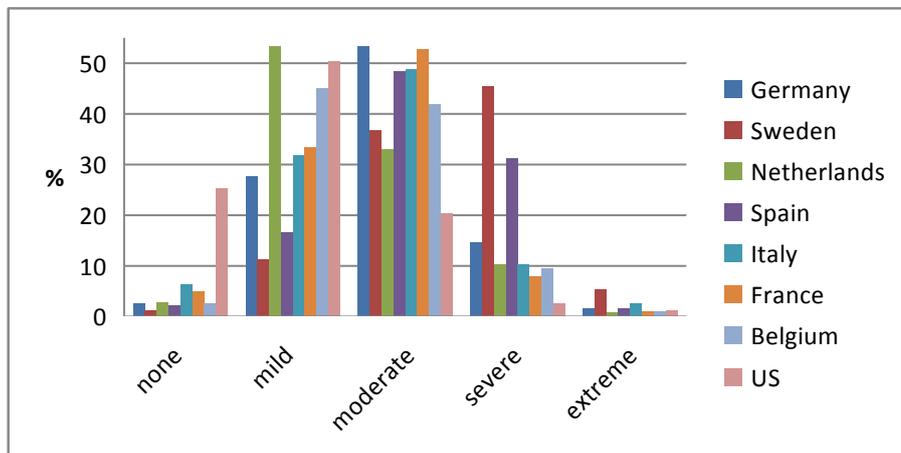
For each vignette, the respondent is asked: “How much is he or she limited in the kind or amount of work he or she could do?” The answer follows a five-point scale: (1) None; (2) Mild; (3) Moderate; (4) Severe; and (5) Extreme/Cannot do any work. Preceding the vignette questions, respondents are asked about their own work limitations: “Do you have any impairment or health problem that limits the kind or amount of work you can do?” with the same answer categories.

Using vignette data, we can identify the reporting heterogeneity for a standardized disability scenario across countries. [Figure 1](#) uses one vignette (Pain Vignette No. 2) as an example to show the different reporting styles for an identical disability scenario. This vignette reads:

*[Name] suffers from back pain that causes stiffness in her back especially at work but is relieved with low doses of medication. She does not have any pains other than this generalized discomfort.*

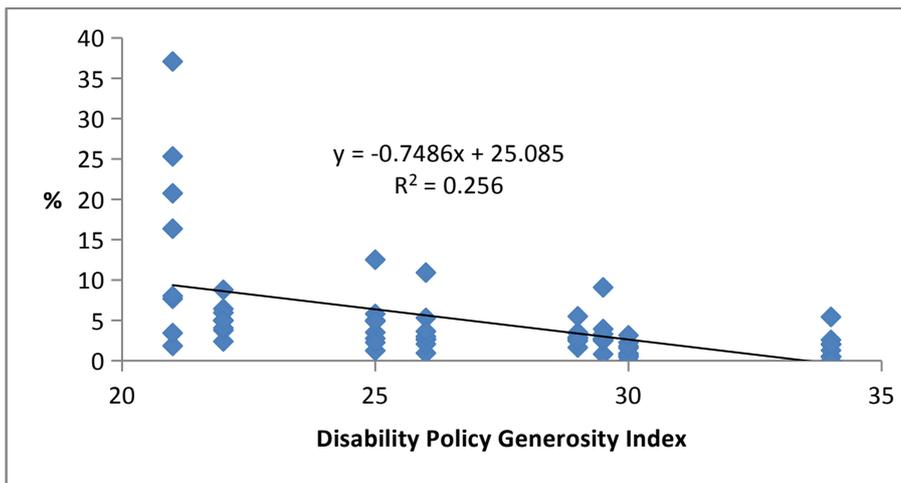
The contrast between countries is striking in how respondents classify the same vignette: More than 25 percent of Americans, compared to only less than 10 percent of Europeans, rate this vignette as not disabled at all. At the other end of the spectrum, as high as 45 percent of Swedes rate this vignette as severely disabled while only 3 percent of Americans rate so.

In [Figures 2-1](#) and [2-2](#), pooling data on all the nine vignettes together, we observed considerable differences in how residents across countries characterize the work

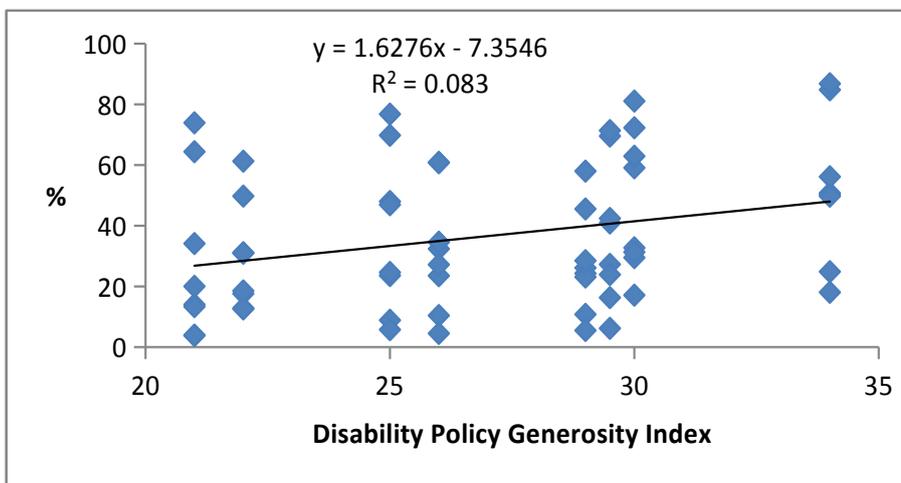


**Figure 1.** Classification for Pain Vignette No. 2 by country

Note: The horizontal axis represents the reported severity of the pain vignette ranging from “none” to “extreme”. The vertical axis represents the percent of respondents who classify the Pain Vignette No. 2 into a certain severity level.



**Figure 2-1.** Disability policy generosity and percentage of respondents classifying 9 vignettes as not at all limited



**Figure 2-2.** Disability policy generosity and percentage of respondents classifying 9 vignettes as severely/extremely limited

limitation severity in these vignettes. European respondents, compared to Americans, tend to classify a vignette as more work-limiting. American respondents are the least generous overall in rating work limitations, followed by the Italians, the Belgians, the Dutch, and the French. The Swedes and the Spaniards are at the other end of the spectrum and inclined to rate a given work limitation as more severe. German respondents are in the middle.

These patterns are notable in light of our discussion of disability program generosity. More inclusive rating scales are associated with more generous disability regimes, as indicated by the fact that the country ranking according to the inclusiveness of disability rating is in line with the ranking in terms of the generosity of the country's disability system.

Figure 2-1 shows the correlation between the generosity of disability system and disability vignette rating. The horizontal axis represents a country's disability policy generosity index. The vertical axis refers to the percentage of respondents in a country who classify a vignette as not at all limited. Each diamond in the graph represents the rating for a specific vignette. There are apparent variations between countries with different disability policy generosity in classifying the severity of the same disability vignettes. The difference between the U.S. and other European countries in rating styles is striking. The graph indicates a weak and negative correlation between disability policy and disability ratings. That is, more generous disability policy seems to be generally associated with more inclusive disability classifying styles. Together with Figure 2-2, the direction of the correlations shows that respondents under more generous disability regimes are more likely to report the same vignettes as more work limiting. The weak strength of the correlations may have to do with how the demographic and socioeconomic characteristics of a population are related to the perception of popular views. Next we will estimate the effects of disability policy generosity on reporting heterogeneity while controlling for a series of respondents' characteristics. We test whether disability policy generosity predicts reporting styles and whether the predictive power is robust to including the respondents' individual level factors.

### 2.3 Covariates

We test a model with a detailed set of individual-level and country-level factors. Specifically, the model includes standard demographic covariates: age dummies, education (in years), dummies for being female, and a series of health indicators (dummies for high blood pressure, diabetes, cancer, lung problems, heart problems, arthritis, obesity, the number of ADL limitations, and the number of IADL limitations). We also include dummies for (last) occupation (technical/sales/administrative support, service, farming/fishing/forestry, precision production/craft/repair, operators/fabricators/laborers, elementary occupation, and managerial/professional specialty as omitted/reference occupation). Past occupation is expected to be an important predictor of a person's risk of having a work limitation. The severity of work limitations is a function not only of the health problems but also of the type of work engaged.

With respect to variation at the country-level, we include a set of variables describing how a country scores in terms of disability policy generosity (coverage and maximum benefit level, disability level for full benefits, permanence of benefits, medical assessment, vocational assessment, minimum disability level and sickness benefit, and unemployment benefit level and duration).

### 2.4 Statistical Approach

Standard ordered regression models (e.g. ordered probit) are often used to analyze self-reported work limitation on a five-point severity scale:

$$\Pr(h_i^s = k) = \Pr(\mu^{k-1} \leq x_i\beta + \varepsilon_i^s < \mu^k) \quad (1)$$

where  $h_i^s$  is respondent  $i$ 's self-reported severity of work limitation,  $k$  is the severity level ranging from 1=*none* to 5=*extreme*,  $x_i$  is a vector of observed characteristics for respondent  $i$ , and  $\varepsilon_i^s$  is a random error term.  $\mu^k$  refers to the cut-point between severity categories on the reporting scale.

The model estimates the probability of a respondent reporting a given category of severity for their work limitation along a latent (continuous) index of work limitation severity that is a function of their individual characteristics. The model also estimates the cut-points,  $\mu^1$  to  $\mu^4$ , as the model parameters, representing the thresholds at which a respondent would change their work limitation ratings along the latent index.

The model assumes that cut-points are constant across individuals. That is, the locations of the cut-points are invariant across respondents (i.e., reporting homogeneity). If this assumption does not hold, in particular, if the cut-points vary with the respondents' characteristic  $x_i$ , then imposing this assumption will lead to biased estimates of the coefficient  $\beta$ . This is because  $\beta$  will reflect both health effects (effects of covariate  $x_i$  on work limitation severity) and reporting effects (effects of covariate  $x_i$  on the cut-points).

To test and estimate flexible models that allow the cut-points to vary with respondents' characteristics, we will need external information to identify the parameters in the cut-points equations. Vignette data can be used as such external information to model the cut-points as functions of respondent characteristics. Respondents are asked to rate identical vignette characters' work limitations. The vignettes are fixed, so the variations in ratings represent differences in response scales used by respondents.

In a generalized ordered probit model, we estimate respondents' severity rating of the vignette character's work limitation:

$$\Pr(h_i^v = k) = \Pr(\mu^{k-1} \leq a^v + \varepsilon^v < \mu_i^k)$$

where

$$\mu_i^k = \gamma_0^k + x_i \gamma^k + z_c \delta^k + c \eta^k \quad (2)$$

Each vignette,  $a^v$ , is constant, plus a random error,  $\varepsilon^v$ . The model estimates the probability of a respondent reporting a given severity category for the vignette character's work limitation, allowing the location of cut-points to vary by respondents' characteristics. Specifically, the vignette rating is estimated as a function of a vignette dummy with each cut-point separately estimated as a function of respondents' characteristics which includes both individual-level factors and country-level factors.

The threshold equation (2) is estimated separately for each of the four cut-points,  $\mu^1$  to  $\mu^4$ , which is determined by respondent  $i$ 's characteristic  $x_i$ , disability policy generosity of country  $c$  in which the respondent resides,  $z_c$ , and other country-specific factors summarized in a vector of country indicator,  $c$ . Our main interest, the effect of disability generosity,  $\delta^k$ , reflects the shift of the cut-point  $\mu^k$  as the disability policy generosity score varies. A negative estimate of  $\delta^k$  would suggest that respondents under a more generous disability regime apply a lower threshold to classify the disability severity level  $k$ , that is, they are more likely to evaluate a given vignette person as more severely work limited. The estimated coefficients for the four cut-point equations are presented in [Table 4](#).

Finally, we impose the cut-points, which are estimated based on respondents' ratings of the vignette characters' work limitations, on the model that estimates respondents' ratings of their own work limitations (equation (1)). The two models, equations (1) and (2), are jointly estimated with the Hierarchical Ordered Probit (HOPIT) procedure suggested by King *et al.* (2004). The vignette model, equation (2), estimates the four cut-points as functions of respondents' characteristics, thus allowing for reporting heterogeneity. The self-reported work limitation model, equation (1), represents the relationship between the respondents' own work limitation severity and their characteristics, with reporting cut-points determined by the vignette model. The

vignette data and approach makes it possible to separately identify health effects and reporting effects rather than a mixture. We illustrate this set of results with counterfactual policy simulations in the Results section. Specifically we focus on the disability policy effects on the reporting scale which in turn shifts the distribution of self-reported disability in the country.

### 3 Results

#### 3.1 Sample Description

The descriptive statistics by country for our analytical sample are provided in [Table 3](#). The table shows large differences in years of education, with low means in the southern European countries. There are also obvious differences in the age composition, with, for example, relatively few 66–70 year olds in Sweden. Most chronic conditions are much more prevalent in the U.S. than in European countries. Still, the distribution of self-reported severity of work disability in the U.S. is quite similar to what is observed in European countries overall (columns “U.S.” vs. “Europe”). However, there are notable differences in the raw distributions across the seven European populations. For example, while all distributions are fairly right-skewed, Sweden are particularly concentrated at “none” and the Netherlands is very concentrated at “none” and “mild”, while the distribution is relatively more equal across the five categories in Belgium.

#### 3.2 Predicting Reporting Scales: Results from Regression Analysis

Tables 4 provides the regression estimates of the respondents’ reporting scales using their ratings on the disability vignettes ( $n=6,652$ ). The estimated effects from four cut-point equations are listed from left to right for cut-point 1 (“not at all limited” to “mildly limited”) to cut-point 4 (“severely limited” to “extremely limited”). As shown in equation (2), the left hand side of each cut-point equation is the location of the cut-point on the severity spectrum. Controls include a detailed set of individual-level and country-level factors. The four cut-point equations are estimated jointly as the four cut-points together determine an individual’s response scale.

In [Table 4](#), a negative coefficient suggests that the respondents apply a lower cut-point when determining the severity level of the work limitation, indicating a more generous reporting style. The results are consistent with systematic reporting heterogeneity at the individual and country level as some of the covariates are found to be predictive of the location of the cut-points. We are particularly interested in, and we show the results on, how the variation in disability policy generosity across countries predicts people’s disability reporting scales.

Some of the policy dimensions are highly correlated, as evidenced by the magnitudes of the correlation coefficients between them. The policies, such as minimum disability level, sickness benefit level and duration, and unemployment benefit level and duration (relative to disability benefits), are strongly and positively correlated, as the correlation coefficients between any two of those dimensions are about 0.7 or more. Moreover, these policy dimensions affect the vignettes’ ratings in the similar fashion, as indicated by the correlation coefficients between each policy dimension and the vignettes classifying. In addition, policy dimensions such as coverage and maximum benefit level are also highly correlated. Therefore, in our regression estimation, we group the correlated policy dimensions to address the potential collinearity.

As shown in [Table 4](#), more extensive coverage and higher maximum benefit level predict more inclusive rating styles at all four cut-points, that is, over the whole work limitation severity spectrum. Permanence of the benefits is also associated with applying lower thresholds, with the strongest effects at the middle of the work limitation spectrum. Easier entry into disability programs with a mild work limitation, reflected in lower minimum disability level, more generous sickness benefits, and better disability benefits compared to unemployment benefits, predicts more inclusive rating styles for work limitation severity.

**Table 3.** Sample means by country

	Belgium	France	Germany	Italy	Netherlands	Spain	Sweden	U.S.	Europe
<i>Demographics and Education</i>									
Female	0.55	0.57	0.56	0.56	0.51	0.56	0.52	0.57	0.55
Age	63.9	64.9	63.8	63.7	62.8	64.8	64.1	64.6	64.1
Age 50–55	0.23	0.22	0.22	0.20	0.24	0.22	0.21	0.20	0.22
Age 56–60	0.21	0.20	0.17	0.23	0.25	0.17	0.20	0.12	0.20
Age 61–65	0.15	0.15	0.21	0.18	0.16	0.15	0.21	0.19	0.17
Age 66–70	0.15	0.13	0.20	0.16	0.14	0.16	0.09	0.23	0.15
Age 70+	0.27	0.30	0.21	0.23	0.21	0.30	0.28	0.25	0.26
Years of education	10.2	8.3	13.1	7.2	11.5	7.1	10.4	12.7	9.6
<i>Health</i>									
High blood pressure	0.27	0.28	0.33	0.35	0.24	0.34	0.30	0.52	0.30
Diabetes	0.08	0.09	0.11	0.09	0.07	0.12	0.07	0.18	0.09
Cancer	0.06	0.06	0.07	0.04	0.05	0.05	0.09	0.12	0.06
Lung problems	0.05	0.06	0.05	0.09	0.06	0.05	0.03	0.09	0.06
Heart conditions	0.13	0.14	0.11	0.09	0.11	0.11	0.12	0.20	0.12
Arthritis	0.24	0.30	0.13	0.34	0.09	0.27	0.09	0.54	0.22
Number of ADL limitations	0.17	0.16	0.09	0.17	0.09	0.13	0.09	0.22	0.13
Number of IADL limitations	0.09	0.12	0.06	0.14	0.07	0.15	0.05	0.16	0.10
Obesity	0.18	0.16	0.16	0.16	0.13	0.23	0.15	0.30	0.17
Self-reported work limitation									
None	0.38	0.51	0.42	0.47	0.54	0.45	0.56	0.45	0.48
Mild	0.36	0.22	0.29	0.29	0.31	0.22	0.16	0.23	0.26
Moderate	0.18	0.18	0.21	0.14	0.09	0.19	0.13	0.18	0.16
Severe	0.07	0.06	0.07	0.06	0.04	0.11	0.11	0.08	0.07
Extreme	0.02	0.02	0.01	0.04	0.02	0.02	0.04	0.06	0.02
<i>Occupation</i>									
Managerial/Professional specialty	0.23	0.18	0.18	0.16	0.24	0.10	0.28	0.27	0.19
Technical/Sales/Administrative support	0.23	0.32	0.28	0.15	0.23	0.09	0.27	0.22	0.23
Service	0.08	0.11	0.12	0.05	0.16	0.07	0.15	0.12	0.11
Farming/Fishing/Forestry	0.02	0.06	0.03	0.06	0.01	0.07	0.03	0.02	0.04
Precision production/Craft/Repair	0.12	0.12	0.13	0.16	0.08	0.13	0.10	0.10	0.12
Operators/Fabricators/Laborers	0.05	0.05	0.08	0.10	0.06	0.10	0.08	0.13	0.07
Elementary occupation	0.13	0.10	0.11	0.11	0.12	0.16	0.06	0.02	0.11
Occupation info missing	0.14	0.06	0.08	0.21	0.10	0.29	0.02	0.12	0.12
<i>No. of observations</i>	543	833	489	426	508	430	402	3021	3631

A more lenient medical assessment regime, meaning that the disability system places greater weight on the opinion of the applicant's treating doctor, seems to predict lower cut-points in rating work limitation severity among its residents. This inclusive style applies to the middle and the right end of the work limitation distribution (cut-points 2, 3 and 4).

The effects of vocational assessment policy are large with the effects concentrated in the first three thresholds. Under a disability regime with relatively lenient vocational assessment, that is, eligibility for disability benefits is based on inability to do one's usual occupation (rather than any job available), a respondent seems to be more likely to classify a given health problem as work limiting except when it comes to rating extremely severe health problems.

**Table 4.** Estimated coefficients on disability policy generosity (multiple policy dimensions) in generalized ordered probit model of respondents' rating of vignette characters' work disability

	Thresholds Equation (Equation (2))			
	Cut-point 1	Cut-point 2	Cut-point 3	Cut-point 4
	Not at all limited $\geq$ Mildly limited	Mildly limited $\geq$ Moderately limited	Moderately limited $\geq$ Severely limited	Severely limited $\geq$ Extremely limited
<i>Disability Policy</i>				
Coverage & Max. benefit level	-0.053***	-0.033***	-0.042***	-0.037***
Disability levels for full disability	-0.040*	0.124***	0.145***	0.030*
Permanence of benefits	-0.072***	-0.115***	-0.141***	-0.030***
Medical assessment	0.084***	-0.127***	-0.179***	-0.040***
Vocational assessment	-0.154***	-0.192***	-0.143***	0.034*
Min. disability level & Sickness benefit & Unemployment benefit	-0.058***	-0.075***	-0.070***	-0.031***
<i>Demographics and Education</i>				
Female	0.035**	0.091***	0.081***	0.029
Age 56–60	-0.013	-0.001	0.025	0.02
Age 61–65	-0.021	0.029	0.074***	0.153***
Age 66–70	-0.006	-0.006	0.034*	0.096***
Age 70+	-0.056**	-0.039**	0.066***	0.224***
Years of education	-0.026***	-0.006***	0.001	0.013***
<i>Health</i>				
High blood pressure	-0.047***	-0.024*	-0.001	-0.009
Diabetes	0.001	-0.034*	-0.055***	-0.111***
Cancer	0.025	0.024	0.049**	0.122***
Lung problems	0.017	-0.023	0.013	0.042
Heart conditions	-0.063***	-0.038**	-0.022	0.02
Arthritis	-0.049***	-0.01	0.006	0.050**
Number of ADL limitation	-0.057***	-0.032***	-0.026**	-0.107***
Number of IADL limitation	0.011	0.005	-0.028**	-0.032*
Obese	-0.002	-0.037**	-0.064***	-0.102***
<i>Occupation</i>				
<i>Ref: Managerial/Professional specialty</i>				
Technical/Sales/Administrative support	-0.029	-0.058***	-0.057***	-0.031
Service	0.021	-0.086***	-0.105***	-0.213***
Farming/Fishing/Forestry	0.048	0.084**	-0.033	-0.083
Precision production/Craft/Repair	-0.068**	-0.03	-0.051**	-0.088***
Operators/Fabricators/Laborers	-0.033	-0.033	-0.106***	-0.189***
Elementary occupation	-0.075*	-0.070**	-0.069***	-0.149***
Constant	-0.780***	0.598***	1.664***	1.576***

Notes: In this estimation, we pool the responses from all the countries together and the total number of observations is 6,652. We included in the estimation an indicator for missing occupation information. \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

Disability levels for full disability benefits show remarkably different effects on reporting styles. More lenient standards to determine disability levels for full disability benefits are associated with stricter rating styles for work limitations. The effects are

concentrated in the middle of the work limitation distribution (cut-points 2 and 3). However, the sign of the effects seems counterintuitive.

While looking at the effects of disability policy generosity and other country-level factor, we also control for a series of respondents' characteristics in the estimation because reporting heterogeneity may be related to differences in population demographics, socioeconomic characteristics, and health distribution in each country. We test whether the predictive power of disability policy generosity is robust to including these respondents' individual-level factors.

Our model (in Table 4) controls for standard respondents' demographic covariates: gender, age (in groups), education (in years), and a series of health indicators (high blood pressure, diabetes, cancer, lung problems, heart problems, arthritis, obesity, and number of Activity of Daily Living or ADL limitations, and number of Instrumental Activity of Daily Living or IADL limitations). Work disability is not purely a health indicator but also depends on the work context. So we also include seven dummies for occupation (technical/sales/administrative support, service, farming/fishing/forestry, precision production/craft/repair, operators/fabricators/laborers, elementary occupation, and managerial/professional specialty as omitted/reference occupation). Some of these individual-level factors significantly predict vignette rating styles, but they do not reduce the predictive power of disability policy generosity.

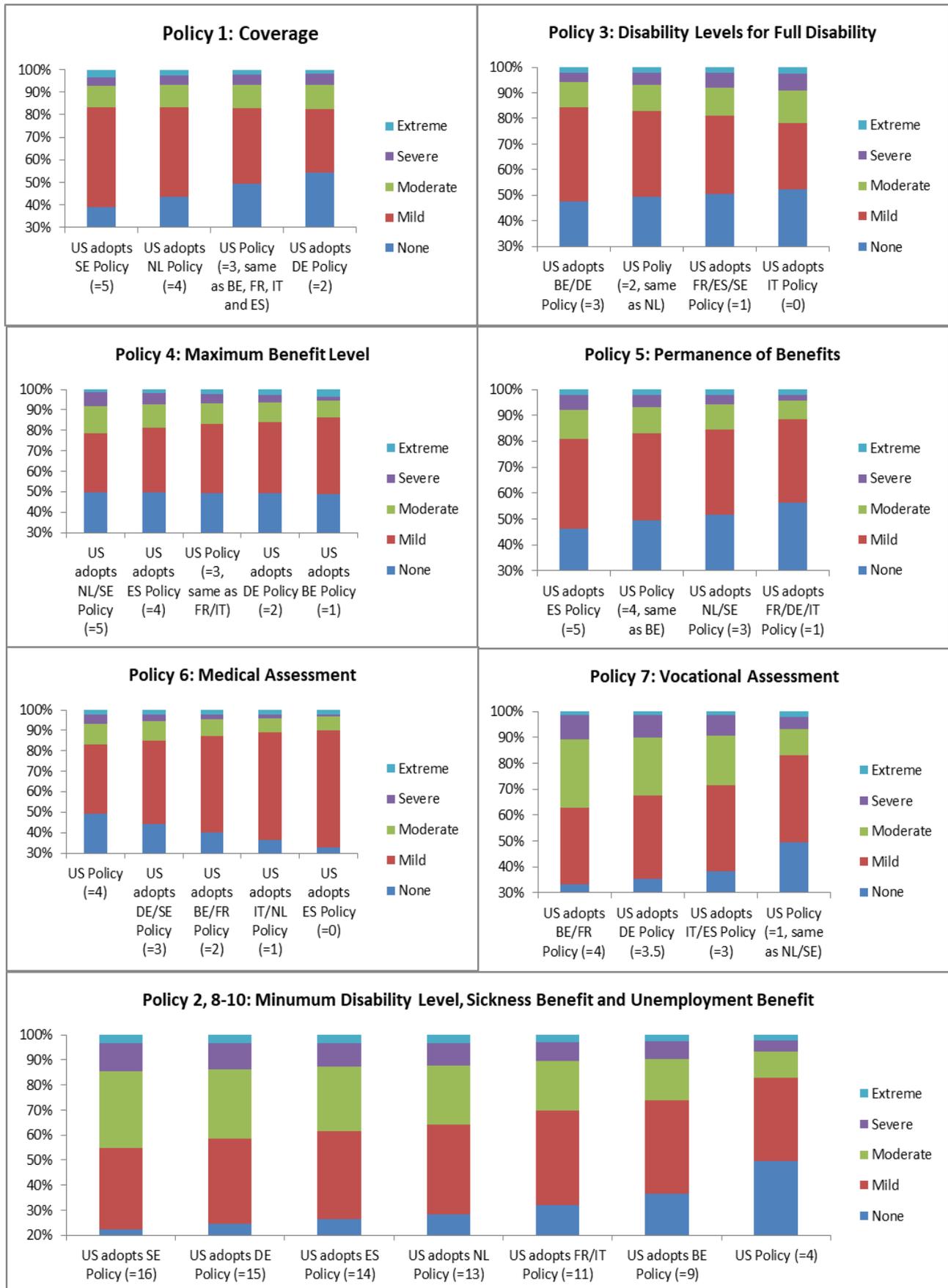
### 3.3 Counterfactual Simulations of the Effects of Disability Policy Generosity

Having estimated the cut-points adjusted for country-level and individual-level factors, we impose the estimated response scales on the model for self-reported work disability. We want to see whether allowing reporting heterogeneity affects the distribution of self-reported work disability. Specifically we focus on the disability policy effects on the response scale which in turn shifts the distribution of self-reported disability in the country. We illustrate such an impact with policy simulations and demonstrate whether changing a country's disability institutional environment would affect what the population's (self-reported) disability distribution looks like. The results are presented in Figure 3. In particular, we show how the disability severity distribution in the U.S. would look like if the U.S. were to adopt other European countries' disability policy. We carry out such simulations for every policy dimension, except for policies 2 and 8–10 which we combine since they are so highly correlated and reflect a country's overall policy towards milder work limitations. All the disability severity distributions in the graphs are adjusted by demographics, health conditions and occupations in the U.S. The exercise performed here is more of a comparative analysis than a simulation. We do not consider any feed-back effect if different policies could change people's behavior.

The first graph depicts the policy simulation results for the first policy dimension, disability policy coverage. The U.S. disability programs cover the labor force plus a means-tested program. The U.S. programs score 3 in terms of their generosity according to OECD and rank in the middle compared to the other countries. If the U.S. were to adopt Sweden's coverage policy, which is the most generous, and covers not only the labor force but the whole population, the simulation shows significant increase in the likelihood of Americans reporting disability. The most pronounced change is an increase of more than 10 percentage points (from 33.5 percent to 44.2 percent) in the proportion of Americans reporting mild work disability, and a drop of more than 10 percentage points in the proportion reporting no disability.

The reporting pattern changes in a similar fashion except with smaller magnitude if we applied the more generous coverage in the Netherlands, where more than the labor force is covered but not the whole population, as compared to the U.S. Tightening the U.S. disability policy to cover only the part of the labor force that is self-insured, similar to the German policy, would make fewer people covered and possibly less likely to characterize their health conditions as work disabling, at least according to the coverage standard of the public disability programs. As shown in the graph, about 5

Figure 3. Predicted disability severity distribution by country



percent of the American population would shift their disability reporting from mild to none if the U.S. adopted the German disability coverage policy.

The second graph presents the simulation results for the policy dimension of disability levels for full benefits. In general, lower disability threshold for full benefits seems to be associated with greater likelihood of reporting work disability and the effect is concentrated in the mild disability category. However, the overall effect is only moderate, which is consistent with what the OECD (2003) finds: although different levels of work incapacity are required for full benefits, the proportions of people receiving full benefits are very similar across countries. It suggests that the specified incapacity levels are just a reference point for people to evaluate the relatively severe work limitations, and the variations in the specific incapacity levels between countries do not seem to matter much.

The third graph shows that changing the maximum disability benefit level would mainly affect the rate of reporting mild and moderate/severe work limitations. On average, a one-point increase in the generosity of maximum disability benefit level is associated with 1 to 2.5 percentage point drop in self-reported mild disability, and about 2 percentage point upswing in the rate of reporting moderate or severe disability. The overall impact seems to be only modest of changing the maximum disability benefit level on the way people rate work limitations.

In the fourth graph, we find that more permanent benefits shift the distribution of the self-assessed disability severity towards the more severe end. The most distinct change appears in the proportion of the population reporting no disability, which reduces by 2 to 3 percentage points for a one-unit increase in the generosity score for the more permanent benefits. Apparently, the permanence of the benefits has only limited effects on disability reporting.

In the fifth graph, we observe that more lenient medical assessment is associated with higher rate of reporting moderate and severe work disabilities, but it is also associated with higher rate of reporting no work disability and much lower rate of reporting mild work disability. A medical assessment is considered more lenient if the process places more weight on the treating doctor's opinion and less likely involves insurance doctor. Every one-unit increase in the generosity score of the medical assessment would lead to a 1.3- to 2.5- percentage point increase in reporting moderate or severe disability, and at the same time a 3.5- to 5- percentage point increase in reporting no disability. This is interesting and may reflect heterogeneous effects in the population: On the one hand, generous medical assessments in the country's public disability evaluation may make some people more likely to classify a given health problem as work limiting; on the other hand, the more relaxed medical assessment in the disability system may be viewed as excessive by others, making them reluctant to rate a given health problem as disabling.

In the sixth graph, more relaxed vocational assessment is associated with more reporting of moderate and severe disabilities and less reporting of no disability. A one-unit increase in the generosity score of vocational assessment would raise the proportion reporting moderate disabilities by 4.5 to 7.8 percentage points and the proportion reporting severe disabilities by 1 to 2 percentage points, and reduce the rate of reporting no disability by 4.3 to 5.7 percentage points.

A vocational assessment is considered stricter if all the available jobs in the national economy, not only one's own previous occupation, are considered in evaluating a disability applicant's residual work capacity. The vocational assessment in the U.S. is strict: to qualify for the federal disability benefits, an applicant has to be deemed unable to do any jobs available in the national economy. In other words, only individuals who do not have residual work capacity would qualify for benefits. In most European countries, the disability system uses one's previous earnings or occupations as reference when determining the individual's eligibility for benefits, especially for partial benefits.

In the simulation, we relax the strict U.S. vocational standard to allow the applicants to engage in a different type of work (likely less demanding) than their previous

occupation and to still qualify to receive benefits. It perhaps would foster a culture of continued attachment to the labor force among people with work limitations in the U.S. and then implant among the general public an idea that work and disability are not necessarily mutually exclusive. In time, that idea would likely affect the way that the public perceives and classifies disability. For example, the general public, who used to call milder health problems “no disability”, would now define them as “mild” or “moderate” disability, as suggested by the simulation result.

In the last graph, we pool together several policy dimensions (minimum disability level, sickness benefit and unemployment benefit) because they are highly correlated and represent the eligibilities for less severe disability cases. Depending on the specific rules in a country, a worker struck by milder work limitations could apply for unemployment benefits, relatively short-term sickness benefits, or some lower level of disability benefits from the public disability program. Our simulation shows that a one-unit increase in the generosity score summarizing these policy aspects, that is, in terms of the systems awarding lower level of disability, would increase the rate of reporting moderate disabilities by 1.3 to 3.1 percentage points and lower the rate of reporting no disability by 1.8 to 2.6 percentage points.

#### 4 Discussion

Disability programs are a substantial and rising component of public social expenditures and an important dimension of the social safety net in many developed countries. A large literature has focused on estimating the effects of disability policy generosity on people’s behavior, such as labor market participation and disability benefit claiming. In this paper, we try to understand how differences in the disability institutional arrangements affects the general public’s views about work disability, an aspect that has received little attention before in disability policy studies.

We first use an anchoring vignette approach to study the role of response scale heterogeneity. We find evidence that, compared to their European counterparts, Americans apply a less inclusive scale to their assessment of work limitations, i.e., they are less inclined to see a given condition as work limiting. The results are consistent with Kapteyn *et al.* (2007) who found Dutch respondents have lower thresholds in reporting work disability than American respondents. We further explore the possible mechanisms for disability reporting heterogeneity among countries. We show that the different reporting styles could be linked to the generosity of disability policies. For example, the less inclusive disability rating styles among Americans than Europeans are associated with the stricter disability policies in the U.S. compared to the European countries. Previous research that utilizes vignette data to study disability reporting heterogeneity has paid little attention to understanding the underlying mechanisms of the observed reporting differences across countries. The two exceptions that have paid attention to the mechanisms either find insignificant effects (Angelini, Cavapozzi and Paccagnella (2012) who have attributed reporting differences to variations in public disability expenditures across countries) or do not formally model the mechanisms (Kapteyn, Smith and van Soest (2009) who have linked reporting differences to work norms across countries).

Our counterfactual policy simulation results suggest an overall positive correlation between disability policy generosity and disability reporting. That is, more lenient disability policies are associated with higher likelihood of reporting disabilities. Different policy dimensions affect the disability classifications in different ways, and the most influential policy dimensions in affecting disability reporting are the policy coverage, medical assessment, and vocational assessment. Specifically, more extensive policy coverage is associated with significantly more reporting of mild disability and less reporting of no disability. More lenient vocational assessments are predicted to result in substantial increases in reporting of moderate and severe disabilities and much less reporting of no disability. Medical assessment does not have such monotonic effects over the disability severity distribution. More relaxed medical assessment

appears to relate to a modest upswing in reporting moderate/severe disabilities and, at the same time, to a much higher rate of reporting no disability.

These findings have important policy implications. For example, when the medical or vocational assessment procedures are changed, in addition to directly affecting individuals' medical and vocational qualifications for disability benefits, over time it would also possibly affect how the general public assesses work disabilities medically and vocationally, especially as time goes by and the policy changes gradually shape the disability culture in the country.

More work-oriented disability policy will likely foster a culture of work in the wake of health limitations. This culture could then quickly evolve over time: when few workers with health impairments take up the work incentives in the disability policy, information is scarce and participation rises slowly. As information accumulates, the effect of employment among the health impaired individuals becomes less uncertain and the participation rate increases.

Here we focused on how longstanding policies and institutions may influence cultural beliefs, rather than the reverse relationship. Culture and institutions are likely interrelated in a complex way. A country shares specific cultural values, such as attitude towards work, sense of solidarity, and preference for redistribution, which may lead to the emergence of particular disability institutions. Then, in turn, certain disability institutions will lead to the survival of certain cultural values and affect the social norms towards work disability. Individuals acquire information about the institutions through social learning, including learning about the policies and assessing the policy results over time.

Given the two-way relationship between policy and beliefs, the effects that we estimate from disability policies to disability perceptions are likely an upper bound of the policy effects. We are not able to analyze the co-evolution of disability policies and disability vignette ratings. For institutions to transform cultural values, it could take a very long time. The main disability institutions in the countries under study here have remained mostly unchanged. Hence, it is plausible to think that the social norms about work disability in those countries have gradually come into being under the influence of the consistent disability institutions in the country. In the future, as data become available on the evolution of both institutions and cultural values, we hope that researchers will revisit these questions to more fully understand the complementarities between disability culture and disability institutions and the dynamic effects of disability reform on disability culture.

## **5 Conclusions**

In this paper, we provide a comparative analysis of disability policy in the U.S. and seven European countries and using an anchoring vignette approach, we investigate whether different views held by the general public regarding what constitutes a work disability are related to cross-country differences in disability policies. We find evidence consistent with the correlation between disability policy generosity and public perceptions about work limitations. A closer look at the correlations reveal that the way people classify disability does not correlate with each policy dimension in the same fashion. The most influential policy dimensions in affecting disability reporting are policy coverage, medical assessment, and vocational assessment. Specifically, more extensive policy coverage and more lenient vocational assessments in a country are linked to its residents rating more vignettes cases as relatively severe disabilities, while a country's more lenient medical evaluations are associated with its citizens classifying more vignettes characters as not disabled at all. Our study has demonstrated an important pathway through which a country's disability policy interacts with its general public, an effect that we should keep in mind while designing and reforming disability programs.

## **Conflict of Interest**

No conflict of interest was reported by all authors.

## Availability of Data

Data used in this study are from the University of Michigan Health and Retirement Study (HRS) and the Survey of Health, Ageing and Retirement in Europe (SHARE). They are publicly available at the data websites: <http://hrsonline.isr.umich.edu/> and <http://www.share-project.org/>.

## Authors' Contributions

Na Yin has contributed to the research question formulation, data analysis and programming, and manuscript writing. Frank Heiland has contributed to manuscript revision.

## Acknowledgements

The authors would like to thank Sanders Korenman, Selcuk Eren, the Editor and the anonymous referees for their helpful comments.

## Appendix A

### OECD Countries' Disability Programs: Benefit Structures and Policy Lessons

Many OECD countries have long used partial disability benefits as a way to encourage people to remain in work, or to return to employment. Most countries with such systems, including some Nordic (Finland, Norway, and Sweden), western (Germany, the Netherlands, and Switzerland), and central and eastern European countries (Czech Republic, Greece, and Hungary), and Korea, offer a full benefit to those assessed to be fully unable to work and various degrees of partial benefits consistent with reduced work capacity. Most countries offer one to four levels of partial benefits, while some use a finer grid (e.g. in 5% intervals for work capacity reduction of 50%–94% in Norway). Partial work capacity is defined in different ways across countries, for example, in terms of the number of hours a person is permitted to work (Germany and Sweden) or in relation to the remaining percentage of work/earnings capacity.

Other countries, including France, Poland, Portugal, and Spain, have a quasi-partial benefit for people who are unable to work in their usual occupation and a full benefit only for those unable to work in any occupation. Hence, the capacity threshold is the same for both benefits but the reference is different. The partial benefit is allowed to be supplemented to some extent by earnings from a job in another occupation. Luxembourg has a similar quasi-partial benefit system but the scheme is somewhat more generous where the full benefit eligibility is measured against former occupation and the partial benefit eligibility is gauged with respect to the last workplace.

Not all the OECD countries have partial benefits in their main disability program. Some countries offer a full benefit to individuals who have fully or partially lost work capacity, such as Australia, Austria, Belgium, New Zealand, Mexico, Slovak Republic, and Turkey. These systems award full benefits to individuals with earnings capacity reduced below a certain threshold, such as 50% in Austria and Mexico, 66% in Belgium, and 40% in Slovak Republic and Turkey, or a threshold defined by the number of hours a person can still work, as in Australia and New Zealand.

Other countries that do not have partial benefits in the main disability program, e.g., Canada, Denmark (after 2003), Italy, Japan, and United Kingdom, adopt a very strict disability definition, one similar to that in the U.S. disability benefits system. However, unlike the U.S., these countries usually have a universal sickness program as a precursor to the long-term disability program and also more alternative public programs to match their disability program. For example, while the Italian disability pension scheme only awards full benefits to people totally unable to work (those with 100% total and permanent incapacity to perform any work), there is a means-tested disability allowance that compensates partial work-capacity loss.

Despite the different benefit structures, the OECD countries have all become increasingly convinced that it is important to identify and utilize residual work capacity of people with health limitations. In principle, people with a partially-reduced work capacity should not leave the labor force and should be supported to find or remain in work. This would help ensure social integration, raise these individuals' living standards and maintain effective labor supply in the face of an aging population. However, in reality, the participation rates of disabled people are often low and not increasing even when the employment rates are increasing for the general working-age population in most countries. On the one hand, economic and labor market changes, such as job requirements changes, perhaps play a role in hindering health-limited individuals from remaining in or returning to work. On the other hand, inadequate policies are one important reason for the low employment rates of these partially disabled individuals. The OECD countries' disability systems have predominately focused on people's work disabilities, rather than work capacity, which has made the systems passive in fostering work. Assessment procedures and benefit systems often push disabled people with significant work capacity into long-term benefit dependency. Countries are increasingly aware of this problem. Many have started to change the approach for those with partial work capacity with a goal of promoting their employment and steering them away from benefit dependency and labor market exclusion.

**Appendix Table A1.** Correlation between generosity of each disability policy dimension and vignettes classifying

	Percentage of respondents classifying the vignettes as <i>Not at All Limited</i>	Percentage of respondents classifying the vignettes as <i>Mildly or Moderately Limited</i>	Percentage of respondents classifying the vignettes as <i>Severely or Extremely Limited</i>
Policy 1	-0.145	-0.145	0.164
Policy 2	-0.439	-0.094	0.192
Policy 3	0.060	0.091	-0.096
Policy 4	-0.139	-0.139	0.157
Policy 5	0.054	-0.122	0.095
Policy 6	0.381	-0.067	-0.035
Policy 7	-0.199	0.086	-0.027
Policy 8	-0.272	-0.021	0.086
Policy 9	-0.503	-0.091	0.205
Policy 10	-0.445	-0.156	0.249

## References

- Alesina A and Giuliano P (2013). Culture and Institutions. *NBER Working Paper No. 19750*. Cambridge, MA, USA: National Bureau of Economic Research.
- Alesina A and Glaeser E (2004). *Fighting poverty in the US and Europe: A world of difference*. New York, NY, USA: Oxford University Press.  
<https://doi.org/10.1093/0199267669.001.0001>
- Angelini V, Cavapozzi D and Paccagnella O (2012). Cross-country differentials in work disability reporting among older Europeans. *Social Indicator Research*, 105(2): 211–226.  
<https://doi.org/10.1007/s11205-011-9878-6>
- Bachrach CA (2013). Culture and demography: From reluctant bedfellows to committed partners. *Demography*, 51(1): 3–25.  
<https://doi.org/10.1007/s13524-013-0257-6>
- Bagod'Uva T, O'Donnell O and van Doorslaer E (2008). Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*, 37(6): 1375–1383.  
<https://doi.org/10.1093/ije/dyn146>
- Bisin A and Thierry V (2010). The economics of cultural transmission and socialization. *NBER Working Paper No. 16512*. Cambridge, MA, USA: National Bureau of Economic Research.

- Boyd R and Richerson PJ (2005). *The origin and evolution of culture*. New York, NY, USA: Oxford University Press.
- Duggan M, Singleton P and Song J (2007). Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls. *Journal of Public Economics*, 91(7–8): 1327–1350.  
<https://doi.org/10.1016/j.jpubeco.2006.12.007>
- Giuliano P (2007). Living arrangements in Western Europe: Does cultural origin matter? *Journal of the European Economic Association*, 5(5): 927–952.  
<https://doi.org/10.1162/JEEA.2007.5.5.927>
- Global Burden of Disease Study 2013 Collaborators (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: A systematic analysis for the global burden of disease study 2013. *The Lancet*, 386(9995): 743–800.  
[https://doi.org/10.1016/S0140-6736\(15\)60692-4](https://doi.org/10.1016/S0140-6736(15)60692-4)
- Government Accountability Office (GAO) (2010). Social security reform: Raising the retirement ages would have implications for older workers and SSA disability rolls. *Briefing for the Senate Special Committee on Aging*, GAO-11-125. Washington, DC, US: General Accounting Office.
- Grol-Prokopczyk H, Freese J and Hauser RM (2011). Using anchoring vignettes to assess group differences in general self-rated health. *Journal of Health and Social Behavior*, 52(2): 246–261.  
<https://doi.org/10.1177/0022146510396713>
- Gruber J (2000). Disability insurance benefits and labor supply. *Journal of Political Economy*, 108(6): 1162–1183.  
<https://doi.org/10.1086/317682>
- Guiso L, Sapienza P and Zingales L (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2): 23–48.  
<https://doi.org/10.1257/jep.20.2.23>
- Kapteyn A, Smith JP and van Soest A (2007). Vignettes and self-reported work disability in the US and the Netherlands. *American Economic Review*, 97(1): 461–473.  
<https://doi.org/10.1257/aer.97.1.461>
- (2009). Work disability, work, and justification bias in Europe and the U.S. *RAND Working Paper WR-696*. Santa Monica, CA, USA: RAND Labor and Population.
- King G, Murray CJL, Salomon JA, et al. (2004) Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1): 191–207.  
<https://doi.org/10.1017/S000305540400108X>
- Livermore G, Mamun A, Schimmel J, et al. (2013) Executive summary of the seventh ticket to work evaluation report: Final report. *Mathematica Policy Research*. Washington, DC, USA: Center for Studying Disability Policy.
- Martin LG and Schoeni RF (2014). Trends in disability and related chronic conditions among the forty-and-over population: 1997–2010. *Disability and Health Journal*, 7(1): S4–S14.  
<https://doi.org/10.1016/j.dhjo.2013.06.007>
- Mullen KJ and Staubli S (2016). Disability benefit generosity and labor force withdrawal. *Journal of Public Economics*, 143: 49–63.  
<https://doi.org/10.1016/j.jpubeco.2016.08.007>
- Murray CJL, Özaltin E, Tandon A, et al. (2003) Empirical evaluation of the anchoring vignettes approach in health surveys. In: Murray CJL and Evans DB (editors). *Health Systems Performance Assessment: Debates, Methods and Empiricism*. Geneva, Switzerland: World Health Organization. p. 369–400.
- Organisation for Economic Co-operation and Development (OECD) (2003). *Transforming disability into ability: Policies to promote work and income security for disabled people*. Paris, France: OECD Publishing.
- (2010). *Sickness, disability and work: Breaking the barriers*. Paris, France: OECD Publishing.
- Sadana R, Tandon A, Murray CJL, et al. (2002). Describing population health in six domains: Comparable results from 66 household surveys. *Global Programme on Evidence for Health Policy. Discussion Paper No. 43*. Geneva, Switzerland: World Health Organization.
- Weathers II RR and Hemmeter J (2011). The impact of changing financial work incentives on the earnings of Social Security Disability Insurance (SSDI) beneficiaries. *Journal of Policy Analysis and Management*, 30(4): 708–728.  
<https://doi.org/10.1002/pam.20611>
- Wise DA (editor) (2017). *Social security programs and retirement around the world: The capacity to work at older ages*. Chicago, IL, USA: University of Chicago Press.
- Yin N (2015). Partial benefits in the social security disability insurance program. *The Journal of Risk and Insurance*, 82(2): 463–504.  
<http://doi.org/10.1111/jori.12028>