Volume Title: Investigations in the Economics of Aging

Volume Author/Editor: David A. Wise, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-90313-3; 978-226-90313-2 (cloth)

Volume URL: http://www.nber.org/books/wise11-2

Conference Date: May 5-8, 2011

Publication Date: May 2012

Chapter Title: Self-Reported Disability and Reference Groups

Chapter Authors: Arthur van Soest, Tatiana Andreyeva, Arie Kapteyn, James P. Smith

Chapter URL: http://www.nber.org/chapters/c12441

Chapter pages in book: (p. 237 - 264)

7

Self-Reported Disability and Reference Groups

Arthur van Soest, Tatiana Andreyeva,
Arie Kapteyn, and James P. Smith

7.1 Introduction

In contrast to other social scientists, economists have long adhered to an individualistic notion of behavior, despite early contributions by, for example, Duesenberry (1949) and Veblen (1899). An important modern contribution to the modeling of social interactions is the seminal work of Becker (1974). Although of wider relevance, Becker's work emphasized the interactions among family members, caused by interdependent utilities as well as a common budget constraint. In more recent years, economists have increasingly recognized that individual actions are fundamentally influenced by the attributes and behaviors of those other individuals who form their social networks; see Topa (2001).

The span of behaviors that have been examined in this new research on social interactions has been expanding rapidly and even a very partial list now includes criminal activity (Glaeser, Sacerdote, and Scheinkman 1996, 2000), neighborhood effects on youth behavior (Case and Katz 1991), models of herd- or copycat-like behaviors (Banerjee 1992), peer effects in education (Hanushek et al. 2003; Ginther, Haveman, and Wolfe 2000), agglomeration economies (Audretsch and Feldman 1996), information

Arthur van Soest is professor in the Econometrics Department of Tilburg University, a research advisor at NETSPAR, and an economist at the RAND Corporation. Tatiana Andreyeva is director of economic initiatives at the Rudd Center for Food Policy and Obesity at Yale University. Arie Kapteyn is senior economist and program director of Labor and Population at the RAND Corporation. James P. Smith holds the Distinguished Chair in Labor Markets and Demographic Studies at the RAND Corporation.

This research was funded by the National Institute on Aging. We are grateful to David Cutler for useful comments. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see http://www.nber.org/chapters/c12441.ack.

exchanges in local labor markets (Topa 2001), labor supply (Woittiez and Kapteyn 1998), consumption (Kapteyn et al. 1997; Alessie and Kapteyn 1991), retirement plan choices (Duflo and Saez 2003), spillovers of cash transfers on noneligibles (Angelucci and Giorgi 2009), effects of lottery winnings on the consumption of neighbors (Kuhn et al. 2011), and social learning through neighbors (Bala and Goyal 1998). As these examples illustrate, the type of social interactions studied has moved well beyond the immediate family to much larger circles of friends, neighbors, and like-minded consumers and workers. Various reasons are given for why these types of social interactions matter, including information sharing, demonstration effects, and the formation of tastes and preferences.

Social interactions may also affect what individuals believe to constitute acceptable or normal behavior based on the standards of the subcommunities in which they live and work. In this chapter, we develop a direct test of this using data from a household survey representative of the Dutch population on how respondents evaluate work disability of hypothetical people with some work-related health problem (vignettes). Combining this with self-reports on the number of people receiving disability insurance (DI) benefits among one's friends and acquaintances, we estimate a model describing the influence of DI prevalence in one's reference group on the subjective scale used to report own and others' work disability.

Both the prevalence of DI benefit receipt and self-reported work disability vary substantially across countries; see Haveman and Wolfe (2000) and Bound and Burkhauser (1999). In particular, both are much higher in the Netherlands than in the United States. Bound and Burkhauser (1999) report that in 1995, the number of DI recipients per 1,000 workers in the age group forty-five to fifty-nine was 103 in the United States, compared to 271 in the Netherlands. Kapteyn, Smith, and van Soest (2007) report that in the age bracket fifty-one to sixty-four, self-reported work disability in The Netherlands is about 58 percent higher than in the United States (35.8 percent in the Netherlands against 22.7 percent in the United States). While the higher level of Dutch participation in DI programs is not surprising given higher DI benefits and easier eligibility compared to the United States,[1] greater Dutch prevalence of self-reported work disability is puzzling as the Dutch population appears to be healthier than the American population.[2]

Kapteyn, Smith, and van Soest (2007) investigated to what extent differences in self-reported work disability can be ascribed to differences in

1. See, for instance, Aarts, Burkhauser, and de Jong (1996). In 2004, DI recipients in the Netherlands made up 13 percent of the labor force (see Statistics Netherlands at http://statline.cbs.nl/StatWeb.), while in the United States, DI recipients constituted 4.8 percent of the civilian labor force (see U.S. Bureau of Labor Statistics at ftp://ftp.bls.gov/pub/news.release/History/empsit.01072005.news).

2. This is suggested by the analysis of a broad set of health conditions by Banks et al. (2008).

reporting styles across countries. Exploiting the vignette methodology originally developed by King et al. (2004), Dutch and US respondents were given the same descriptions of work disability problems for hypothetical persons ("vignettes"). Dutch respondents appeared to be much more likely to describe the same work disability problem as constituting a work disability than American respondents. Kapteyn, Smith, and van Soest (2007) found that more than half of the observed difference in self-reported work disability between the two countries can be explained by this difference in response styles.

This result implies that US and Dutch respondents have different norms for evaluating work disability. Our chapter analyzes to what extent this is due to peer group effects: do respondents with many DI recipients in their peer group have social norms that make them more likely to evaluate given health problems as constituting a work disability?

We formalize this notion by introducing the concept of prevalence of DI benefit receipt in one's *reference group,* defined as one's circle of friends and acquaintances. In a Dutch survey that we designed and implemented, we asked respondents directly how many people among their friends and acquaintances receive DI benefits. In this chapter, we develop a model that jointly explains the categorical answer to this question and self-reported work disability. The main feature of the model is the notion that response scales for reporting no, mild, or severe work disability can be affected by a "peer group effect," that is, by the number of people in the reference group receiving disability benefits. To identify the determinants of response scales, we exploit anchoring vignettes as in Kapteyn, Smith, and van Soest (2007).

Using this additional information helps to solve the identification problem that is present in many models with peer group effects, known as the *reflection problem* (Manski 1993). Because our reference group variable refers to DI receipt in the reference group and not perceived disability, it seems reasonable to assume that this variable is uncorrelated to the unobservables driving the individuals' norms. This makes it possible to include reference group DI receipt as an exogenous variable in the vignette evaluations. Because the actual disability of the hypothetical vignette persons is by design independent of any respondent characteristic, the effect of reference group DI on vignette evaluations must be an effect on the respondent's norms used to evaluate (own or the vignette person's) work disability.

The remainder of the chapter is organized as follows. In the next section, we briefly describe the micro-data used in our analysis. Section 7.3 presents the model, which essentially consists of three equations. One equation explains the answers to the question about DI benefit receipt in the respondents' reference group. A second equation models self-reported work disability. The third equation (or rather set of equations) explains how individual response scales to questions on work disability (or anchoring vignettes) are affected by the prevalence of DI benefit receipt in the reference group. Throughout, we control for a large number of other variables, such as sociodemographic characteristics and health conditions.

Section 7.4 summarizes our main results. We find that DI benefit receipt in one's reference group has a significant effect on response scales in the expected direction. To gauge the size of this effect, we graph the relation between DI benefit receipt in the reference group against self-reported work disability. It turns out that to explain the complete difference in response scales between the United States and the Netherlands, the percentage of respondents in The Netherlands reporting to know at least some DI benefit recipients has to fall by about 25 percent. This is an order of magnitude that seems reasonable given the substantial difference in the number of Dutch and US people on DI benefits. The final section presents our conclusions.

7.2 The Data

In this research, we use information obtained from the Dutch CentERpanel. This is an Internet panel of about 2,250 households who have agreed to respond to a survey every weekend. Respondents are recruited by telephone. If they agree to participate and do not already have Internet access, they are provided with Internet access (and, if necessary, with a set-top box that can be used together with their television screen). Thus, the CentERpanel is not restricted to households with Internet access, but representative of the Dutch adult population except the institutionalized. Sample weights based upon data from Statistics Netherlands are used to correct for unit nonresponse. The sample that we use to estimate our model consists of about 2,000 respondents who participated in several interviews with questions on work disability in 2003.

From multiple waves of the data that have been collected in the past, the CentERpanel has a rich set of variables on background characteristics of the respondent and household, including their income and labor market status and several salient dimensions of health. In August 2003, we collected work disability self-reports and vignette evaluations (described in the following). In October 2003, we fielded a second wave of vignettes with slightly different wording of the questions and also included questions about reference groups. For our analysis, we will use the vignette and reference group data from this October wave. Appendix A lists the vignette questions. All vignettes are presented with either a female or a male name.[3]

For each of the vignettes, the respondent is asked the following question:

> *"Does . . . have a health problem that limits the amount or type of work he/she can do?"*

with a five-point response scale:

3. Female or male names are assigned randomly. In appendix A, we only show one of the two names per vignette.

Table 7.1 Frequencies for vignette answers (CentERpanel, October 2003)

Affect vignettes	Affect 1	Affect 2	Affect 3	Affect 4	Affect 5
Not at all limited	41.2	96.2	11.1	18.7	2.2
Somewhat limited	49.7	2.8	44.3	44.8	8.4
Moderately limited	7.2	0.6	31.2	26.0	18.6
Severely limited	1.4	0.5	12.2	8.9	40.4
Extremely limited/cannot work	0.5	0.0	1.3	1.6	30.4
Pain vignettes	Pain 1	Pain 2	Pain 3	Pain 4	Pain 5
Not at all limited	22.5	8.2	0.6	0.3	0.8
Somewhat limited	61.8	47.1	6.6	6.2	12.9
Moderately limited	13.4	34.1	25.7	29.4	31.3
Severely limited	1.9	9.2	49.5	43.2	39.2
Extremely limited/cannot work	0.4	1.4	17.6	20.9	15.9
CVD vignettes	CVD 1	CVD 2	CVD 3	CVD 4	CVD 5
Not at all limited	91.2	10.6	1.8	20.7	6.7
Somewhat limited	7.8	46.2	18.2	44.9	34.1
Moderately limited	0.9	29.2	32.6	25.0	30.3
Severely limited	0.1	11.8	33.6	8.8	20.7
Extremely limited/cannot work	0.0	2.3	13.9	0.6	8.3

Notes: Data are weighted. $N = 1,980$ (complete sample). See appendix A for the wordings of the vignette questions. CVD = cardiovascular disease.

not at all; yes, mildly limited; yes moderately limited; yes, severely limited; yes, extremely limited/cannot work.

Table 7.1 presents the response frequencies for each of the fifteen vignette questions. The differences in distributions of answers correspond quite well with the variation in severity of the conditions described in the vignettes. For example, in all three domains of affect, pain, and CVD (cardiovascular disease), the condition described in the third vignette seems much more severe than that described in the first, and respondents ranked them accordingly. Moreover, there was also a great deal of consistency among respondents in how they ordered vignettes in terms of their severity, showing that respondents understood these experiments and took their responses seriously; see Banks et al. (2008) for details.[4]

Table 7.2 presents the distribution of the answers to the question on own work limitations by age group. These represent answers to the following question:

4. This does not imply that everyone ranks the vignettes in the same order. In some cases where average vignette rankings are quite similar (such as the vignettes Pain 3 and Pain 4 in table 7.1), order reversals frequently occur. The econometric model is able to capture this using idiosyncratic errors in the vignette evaluations; see section 7.3.

Table 7.2 **Distribution of self-reported work disability by age (%)**

	Age group						
	15–24	25–34	35–44	45–54	55–64	65+	Total
Not at all limited	86.8	74.1	69.2	55.9	52.8	48.4	63.1
Somewhat limited	5.4	20.7	17.5	24.2	28.5	34.3	22.8
Moderately limited	5.8	3.2	5.8	7.0	10.5	10.9	7.1
Severely limited	2.0	0	2.1	2.9	1.8	3.7	2.2
Extremely limited/cannot work	0	1.8	5.4	9.9	6.3	2.8	4.8
No. of observations	68	362	438	460	336	316	1,980

Notes: Data are weighted. $N = 1,980$ (complete sample).

"Do you have an impairment or health problem that limits you in the amount or kind of work you can do?".

The question allows respondents to reply on the five-point scale:

(1) No, not at all, (2) Yes, I am somewhat limited, (3) Yes, I am rather limited, (4) Yes, I am severely limited, (5) Yes, I am very severely limited—I am unable to work.

These response categories are identical to the ones used to gauge the severity of the vignette work limitations.

Table 7.2 implies that about 37 percent of the Dutch population reports to have at least a mild work limitation and about 14 percent have a work-limiting health problem or impairment that they gauge as moderately limiting or worse. Not surprisingly, work-related health deteriorates with age (although cohort effects may also play some role in this pattern).

The most interesting groups are probably people in the age groups forty-five to fifty-four and fifty-five to sixty-four. For them, the prevalence of work-limiting health problems is large, and this will often be an important reason not to participate in the labor market. For the sixty-five-plus, work-limiting health problems are even more prevalent, but these people are almost always retired anyhow because the Netherlands has mandatory retirement at age sixty-five for almost all employees.

Appendix B presents some of the questions about reference groups asked in the October wave and used in the empirical analysis. Our operationalization of a reference group is the circle of acquaintances mentioned in these questions. The first two reference group questions provide information on the modal age and modal education level in the respondent's reference group. In the analysis, we will combine the age and education categories into a smaller number of broader brackets. Table 7.3 presents descriptive statistics for our independent variables, including the responses to the first two reference group questions listed in appendix B. For example, 27 percent of

Table 7.3 Sample statistics for independent variables

	Mean or percent
Stroke	1.3
Cancer	3.8
Lung disease	6.0
Heart disease	7.1
High blood pressure	19.2
Diabetes	4.8
Emotional problems	11.0
Arthritis	10.4
Problems with vision	3.8
Often pain	25.4
Age in years	47.6
Low education level	39.1
Medium education level	38.7
High education level	22.1
Female	49.9
Northern provinces	14.3
Eastern provinces	21.6
Western provinces	38.7
Southern provinces	25.5
Age in reference group	
<25	8.7
25–35	20.2
36–45	27.0
46–55	19.7
56–65	14.7
66+	9.8
Low education level in reference group	24.9
Medium education level in reference group	47.9
High education level in reference group	27.2

Notes: Data are weighted. N = 1,764 (estimation sample). All variables other than "Age in years" are dummies. The table gives the percentage of observations for which the dummy has a value of 1. Northern provinces are Groningen, Friesland, and Drenthe; eastern provinces are Overijssel, Flevoland, and Gelderland; western provinces are Utrecht, Noord-Holland, and Zuid-Holland; southern provinces are Zeeland, Noord-Brabant & Limburg.

all respondents report that most of the people in their reference group are in the age group thirty-six to forty-five. About 48 percent say that most of their acquaintances have a medium education level (while almost 39 percent of the respondents has that level).

The other reference group questions refer to the number of acquaintances receiving disability benefits, separately for men and women. These are the crucial variables for our analysis as they measure DI benefit receipt in the reference group. For men, we will use the number of male acquaintances on disability benefits; for women, we will only consider the female acquaintances. We discuss the sensitivity of our results to this definition of the reference group variables in section 7.4.1.

Table 7.4 **Distribution of disability in the reference group by age (%)**

	Age group						
	15–24	25–34	35–44	45–54	55–64	65+	Total
Men							
None	82.9	65.6	52.5	55.1	39.4	53.8	56.7
Very few	17.1	31.5	41.5	36.6	44.1	34.7	35.5
A few/many	0	2.9	5.9	8.4	16.5	11.4	7.8
No. of observations	29	174	221	248	196	199	1,067
Women							
None	76.4	67.8	60.7	62.6	58.9	55.2	62.6
Very few	23.6	29.0	35.7	30.4	32.9	38.2	32.4
A few/many	0	3.2	3.6	7.1	8.2	6.5	5.0
No. of observations	39	188	217	212	140	117	913

Notes: Data are weighted. $N = 1,980$ (complete sample).

The distribution of reported DI receipt in the reference group by gender and age group is presented in table 7.4. Here, and in the rest of the chapter, we combine the categories of prevalence of DI receipt in the reference group to three: "Nobody," "Very Few," "A Few/Many," because the frequencies for "Few" and particularly "Many" are small. Young people typically know no one on disability benefits. The number of reference group members on disability benefits is highest for fifty-five to sixty-four-year-old respondents, who also most commonly receive disability benefits themselves. People older than sixty-five may often have a work disability (see table 7.2), but table 7.4 shows they hardly ever receive disability benefits—they receive a state pension and usually one or more additional occupational pensions. The number of women on disability benefits in women's reference groups is typically smaller than the number of men on disability benefits in men's reference groups, particularly at older ages. This may be because women in older cohorts often stopped working at an early age (usually to raise children) and never qualify for disability benefits after that.

Plausibly, these reference group variables are endogenous to the respondent's own work disability—respondents who have a work disability will often not work and will not only receive disability benefits, but will also more easily get acquainted with other people on disability benefits. Hence, we will treat the number of acquaintances on disability benefits as a dependent variable, modeled jointly with work limitations. Table 7.5 shows cross tabs of self-reported work limitations and self-reported prevalence of DI receipt in one's reference group. For simplicity of presentation, we combine categories for self-reported work disability to three: "Not Limited," "Mildly Limited," "Moderately Limited/Severely Limited/Extremely Limited." The table clearly illustrates a positive relation between self-reported work limita-

Table 7.5 Self-reported work disability and reference group disability

	Disability in the reference group (%)			
Self-reported work disability	None	Very few	A few/many	Total
Not limited	60.4	35.0	4.6	100.0
	70.6	66.2	41.5	66.9
Mildly limited	55.4	34.7	9.9	100.0
	21.7	22.0	30.0	22.4
Moderately, severely, and extremely limited	41.3	39.2	19.6	100.0
	7.7	11.9	28.5	10.7
Total	57.3	35.4	7.4	100.0
	100.0	100.0	100.0	100.0

Notes: Data are weighted. N = 1,764 (estimation sample). Row percentages (first line) and column percentages (second line).

tions and the number of people in one's reference group drawing disability benefits.

There are several competing explanations for this positive association. First of all, there may be a causal effect of the prevalence of DI receipt in one's reference group on the tendency to report work limitations. Second, as discussed in the preceding, it is possible that respondents with work limitations are more likely to associate with others who have a work disability (e.g., because of the existence of networks of people with work disabilities). Third, there may be other (observed or unobserved) factors that both increase the likelihood that respondents have a work limitation and the probability that they know others with work limitations. One such factor is age. Fourth, response scales used in answering the reference group questions might be correlated with response scales in self-reported work disability. Respondents may, for instance, exaggerate the number of friends or acquaintances on DI to "justify" their own report of a work limitation (Bound 1991). These explanations are not mutually exclusive. We think these explanations are the most plausible ones, but undoubtedly there are more. For example, knowing many people on disability benefits might increase genuine work disability. We are particularly interested in the role played by the first explanation, reflecting a social interaction effect. In the next section, we present a model that aims at isolating the importance of the first explanation; in the discussion of the results, we will also return to the competing explanations.

7.3 A Model with Reference Groups

Our econometric model explains the reported number of people receiving disability benefits in the reference group R (see table 7.4), self-reported work disability Y (see table 7.2), and reported work disability of the fifteen vignette persons $Y^1, \ldots, Y^{15}$ (see table 7.1).

7.3.1 Self-Reports of Own Work Disability

Individuals evaluate the extent of their work disability with a self-evaluation of whether their health problems and working conditions are sufficiently problematic to place them above their own subjective threshold of being somewhat limited or more than somewhat limited. The result of that evaluation depends on the extent of their true health problems as well as their subjective thresholds of what constitutes a disability, both of which vary across individuals.

More formally, self-reported work disability Y of respondent i is modeled on a three-point scale of "not at all limited," "somewhat limited," and "more than somewhat limited" (combining the three most serious categories "moderate," "severe," and "extreme," to one) as follows:

$$Y_i^* = X_i\beta + \varepsilon_i \tag{1}$$

$$Y_i = j \text{ if } \tau_i^{j-1} < Y_i^* \leq \tau_i^j, j = 1, 2, 3. \tag{2}$$

For notational convenience, we define $\tau_i^0 = -\infty$ and $\tau_i^3 = \infty$. The remaining thresholds τ_i^1 and τ_i^2 will be modeled as functions of observable and unobservable respondent characteristics as described in section 7.4.1. The error term ε_i is assumed to be standard normally distributed. (Complete assumptions on error terms are given in section 7.3.5.)

Because thresholds depend on respondent characteristics, self-reported work disability alone is not enough to distinguish between variation in Y_i^* (that is, genuine variation in work-related health) and variation in the thresholds (that is, variation in what constitutes a disability in respondents' perceptions). Vignettes are used to identify this distinction.

7.3.2 Vignette Evaluations

The vignettes provide all respondents with the descriptions of the same set of work disability problems. As a consequence, variation in how respondents evaluate the given health problems informs us about variation in the subjective thresholds used by the respondents. More formally, the evaluations Y_i^l of vignettes $l, l = 1, \ldots, 15$, are given by

$$Y_i^{l*} = \theta^l + \delta F_i^l + \varepsilon_i^l \tag{3}$$

$$Y_i^l = j \text{ if } \tau_i^{j-1} < Y_i^{l*} \leq \tau_i^j, j = 1, 2, 3. \tag{4}$$

Here F_i^l is a dummy variable indicating whether the person described in the vignette is female ($F_i^l = 1$) or male ($F_i^l = 0$). This specification follows earlier work by Kapteyn, Smith, and van Soest (2007), who find that respondents (both males and females) tend to be "harsher" on female than on male vignette persons, that is, $\delta < 0$. We assume that all ε_i^l are independent of each other and of the other error terms and follow a normal distribution with mean zero and variance σ_υ^2. Thus, the ε_i^l are interpreted as idiosyncratic noise driving vignette evaluations; they reflect arbitrariness in each separate evaluation. If respondents have a persistent tendency to give low or high evaluations, this will not be captured by ε_i^l but by an unobserved heterogeneity term in the response scales; see section 7.3.3.

7.3.3 Response Scale Thresholds

The crucial assumption guaranteeing that vignettes help to identify response scale differences is that individuals use the same scales in evaluating themselves as they do with the vignette persons (*response consistency;* see King et al. 2004). The thresholds used in the vignette evaluation can vary across all types of individual attributes. In this study, we expand the set of attributes and include the number of persons among friends and acquaintances who are on disability benefits R_i^*. The thresholds τ_i^1 and τ_i^2 are modeled as follows:

(5) $$\tau_i^1 = X_i\gamma_1 + \gamma_1^R R_i^* + \xi_i$$

(6) $$\tau_i^2 = \tau_i^1 + e^{X_i\gamma_2 + \gamma_2^R R_i^*}.$$

We have included the vector X_i of respondent characteristics (independent of all error terms) to allow for a rather general way in which response scales vary with individual characteristics. The distance between the two thresholds is also allowed to depend on these characteristics. The exponential forces it to be positive, as in King et al. (2004). The key parameters of interest are γ_1^R and γ_2^R, the estimated impact of the number of people on DI in one's reference group on the threshold that is used to evaluate work disability. In particular, γ_1^R is expected to be negative: people who know many people on disability benefits will think of work disability as something common and will more often evaluate people (including themselves) as work disabled, thus using lower thresholds.[5]

The term ξ_i reflects unobserved heterogeneity in thresholds. For computational convenience, we do not allow for unobserved heterogeneity in the distance between the two thresholds. ξ_i is assumed to follow a normal distribution with variance σ_ξ^2, independent of X_i and all other unobservables in the model except one: the unobserved component of the thresholds driving the answer to the question how many people in the respondent's reference group receive disability benefits (ϕ_i^1 and ϕ_i^2; see section 7.3.4).

7.3.4 DI Receipt in the Reference Group

As explained in the preceding, we consider DI receipt in the respondent's reference group of the respondent's own sex and combine the outcomes

5. In the empirical work, we will allow the parameters γ_1^R and γ_2^R to depend on education level, age, and gender. For notational convenience, we do not make this explicit in the notation.

"few" and "many" because of the small number of observations with the latter outcome. Thus, we obtain an ordered response variable with three possible outcomes, $j = 1$ ("none"), $j = 2$ ("very few"), and $j = 3$ ("a few" or "many"). This will be modeled with an ordered probit equation:

$$(7) \qquad R_i^* = X_i\beta^R + \omega_i^R,\ \omega_i^R \sim N(0, \sigma_\omega^2)$$

$$(8) \qquad R_i = j \text{ if } \phi_i^{j-1} < R_i^* \leq \phi_i^j,\ j = 1, 2, 3.$$

For notational convenience, we define $\phi_i^0 = -\infty$ and $\phi_i^3 = \infty$. In the following, we will further specify the thresholds ϕ_i^1 and ϕ_i^2. The vector X_i of respondent characteristics driving DI receipt in the reference group is assumed to be independent of all the errors in the model. Equation (7) has a "reduced form" nature in the sense that we do not explicitly model how work disability and labor force status affect disability in the reference group. The exogenous determinants of labor force status and disability are included among the regressors X_i to account for this.

Because it is likely that there are common unobserved factors affecting both the number of people one knows on disability benefits and one's own evaluation of work disability, we allow for a nonzero correlation coefficient ρ between ε_i and ω_i^R. This correlation also allows for the role of actual labor force status (which is not included explicitly in the model but "substituted out"): work disability drives labor force status, and labor force status drives the composition of the reference group.

We allow for a common unobserved heterogeneity component driving the thresholds $\tau_i^j, j = 1, 2$ and the thresholds in the reference group equation ϕ_i^k, $k = 1, 2$ by specifying $\phi_i^1 = \phi_{0,1} + \mu\xi_i$ and $\phi_i^2 = \phi_{0,2} + \mu\xi_i$. We normalize $\phi_{0,1} = 0$. The parameter μ could be positive (respondents exaggerating their work disability also exaggerate the number of their acquaintances on DI) or negative (respondents who think of work disability as something exceptional will tend to interpret a given number of acquaintances on DI as large).[6] $\phi_{0,2}$ and μ are additional parameters to be estimated. Define $u_i^R = \omega_i^R - \mu\xi_i$. By way of normalization, we set $\text{Var}(u_i^R) = 1$. We can then rewrite equation (8) as

$$(9) \qquad R_i = j \text{ if } \phi_0^{j-1} < X_i\beta^R + u_i^R \leq \phi_0^j,\ j = 1, 2, 3.$$

7.3.5 Error Terms and Identification

The error terms in the model, including unobserved heterogeneity components, are ε_i, ε_i^l, $l = 1, \ldots, 15$, ω_i^R, and ξ_i. We assume they are all normally distributed and independent of the regressors X_i and F_i^l. The only correlation we allow for is between ε_i and ω_i^R. We assume $(\varepsilon_i, \omega_i^R)$ is bivariate normal with correlation coefficient ρ. The assumption that ξ_i is independent of ε_i implies

6. It seems natural to add another error term to the ϕ_i^j that is independent of everything else, but this will be subsumed in ω_i^R.

that people with higher thresholds do not tend to have larger or smaller genuine work disability (on a continuous scale), keeping observed characteristics X_i and V_i constant. The assumption seems quite plausible although one might argue that lower thresholds point at unobserved characteristics such as pessimistic views that can also genuinely reduce respondents' ability to work. As we shall see, the assumption is largely innocuous and does not affect identification of the structural parameters. To judge to what extent our assumptions impose restrictions and to investigate identification, it is useful to rewrite the model introduced so far somewhat.

Combine equations (1) and (2) to obtain

$$Y_i = j \text{ if } \tau_i^{j-1} < X_i\beta + \varepsilon_i \leq \tau_i^j, j = 1, 2, 3. \tag{10}$$

Similarly, combine equations (3) and (4):

$$Y_i^l = j \text{ if } \tau_i^{j-1} < \theta^l + \delta F_i^l + \varepsilon_i^l \leq \tau_i^j, j = 1, 2, 3. \tag{11}$$

Combining equation (10) with equations (5) and (6) leads to the following observational rule for observed work disability reports:

$$\begin{aligned} (12)\quad & Y_i = 1 \text{ if } X_i\beta + \varepsilon_i \leq X_i\gamma_1 + \gamma_i^R R_i^* + \xi_i \\ & Y_i = 2 \text{ if } X_i\gamma_1 + \gamma_1^R R_i^* + \xi_i < X_i\beta + \varepsilon_i \leq X_i\gamma_1 + \gamma_1^R R_i^* + \xi_i + e^{X_i\gamma_2 + \gamma_2^R R_i^*} \\ & Y_i = 3 \text{ if } X_i\beta + \varepsilon_i > X_i\gamma_1 + \gamma_1^R R_i^* + \xi_i + e^{X_i\gamma_2 + \gamma_2^R R_i^*} \end{aligned}$$

Inserting equation (7) into equation (12), this can be rewritten as

$$\begin{aligned} (13)\quad & Y_i = 1 \text{ if } X_i[\beta - \gamma_1 - \gamma_1^R\beta^R] \leq \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i \\ & Y_i = 2 \text{ if } \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i < X_i[\beta - \gamma_1 - \gamma_1^R\beta^R] \\ & \qquad \leq \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i + e^{X_i[\gamma_2 + \gamma_2^R\beta^R] + \gamma_2^R\omega_i^R} \\ & Y_i = 3 \text{ if } X_i[\beta - \gamma_1 - \gamma_1^R\beta^R] > \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i + e^{X_i[\gamma_2 + \gamma_2^R\beta^R] + \gamma_2^R\omega_i^R}. \end{aligned}$$

Similarly, combining equation (11) with equations (5) and (6) and inserting equation (7) yields

$$\begin{aligned} (14)\quad & Y_i^l = 1 \text{ if } \theta^l + \delta F_i^l - X_i[\gamma_1 + \gamma_1^R\beta^R] \leq \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i^l \\ & Y_i^l = 2 \text{ if } \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i^l < \theta^l + \delta F_i^l - X_i[\gamma_1 + \gamma_1^R\beta^R] \\ & \qquad \leq \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i^l + e^{X_i[\gamma_2 + \gamma_2^R\beta^R] + \gamma_2^R\omega_i^R} \\ & Y_i^l = 3 \text{ if } \theta^l + \delta F_i^l - X_i[\gamma_1 + \gamma_1^R\beta^R] > \xi_i + \gamma_1^R\omega_i^R - \varepsilon_i^l \\ & \qquad + e^{X_i[\gamma_2 + \gamma_2^R\beta^R] + \gamma_2^R\omega_i^R}. \end{aligned}$$

For completeness, we repeat the equation for reference group disability (equation [8]):

$$R_i = j \text{ if } \phi_0^{j-1} < X_i\beta^R + u_i^R \leq \phi_0^j, j = 1, 2, 3. \tag{15}$$

We can see from equations (13), (14), and (15) that the stochastic behavior of the system is determined by the following composite error terms:

$\xi_i + \gamma_1^R\omega_i^R - \varepsilon_i$, $\xi_i + \gamma_1^R\omega_i^R - \varepsilon_i^l$, u_i^R ($= \omega_i^R - \mu\xi_i$), and ω_i^R (in the exponent). All of these error terms are allowed to be correlated with each other, the only restriction being that the covariance matrix of $\xi_i + \gamma_1^R\omega_i^R - \varepsilon_i^l$, ($l = 1, \ldots 15$) has a one-factor structure.

Next we turn to identification. First consider equation (15). Making the normalizing assumptions that $\phi_0^1 = 0$ and $\text{Var}(u_i^R) = 1$, the vector β^R is identified. The vignette equations (equation [14]) next identify θ^l, δ, $\gamma_1 + \gamma_1^R\beta^R$ and $\gamma_2 + \gamma_2^R\beta^R$, where we normalize $\theta^1 = 0$. Because $\gamma_1 + \gamma_1^R\beta^R$ is identified, β is identified from equation (13). The remaining issue is how to identify γ_1^R and γ_2^R. Because $\gamma_1 + \gamma_1^R\beta^R$, and β^R are identified, we can identify γ_1^R if there is at least one exclusion restriction on γ_1. In other words, equation (15) needs to contain at least one X variable that is not present in equation (5). A similar exclusion restriction identifies γ_2^R. Once γ_1^R and γ_2^R are identified, γ_1 and γ_2 are identified as well.

Thus, identification of the reference group effect requires exclusion restrictions—variables that affect DI receipt in the reference group but do not have a direct effect on the evaluation threshold. For this, we use the directly elicited reference group variables on the typical age and education of respondents' acquaintances. These variables are allowed to affect response scales (represented by the thresholds τ_i^j) only through the reference group variable R_i^*. Because there are more reference group variables than needed for identification, we can perform a test exploiting overidentifying restrictions to investigate the plausibility of the exclusion restrictions. As we will see in the empirical results section, the restrictions are not rejected by the overidentification test.

As in all models with reference group effects, identifying the causal effect of the reference group variable requires model assumptions due to endogeneity issues and confounding effects (cf. Manski 1993). A crucial difference with the case discussed by Manski (1993) is that we have direct information on reference group disability receipt. As we have seen in the preceding, this identifies β^R and, hence, in combination with at least one exclusion restriction, we can identify γ_1^R and γ_2^R.

7.4 Results

We estimate the model using simulated maximum likelihood. Details of the likelihood function are presented in appendix C. The integrals in the likelihood contributions (equation [C6] in appendix C) are replaced by smooth simulation-based approximations, by drawing 200 times from the joint distribution of ξ and u^R and using Halton draws.[7] Experiments with a substantially larger number of draws did not lead to appreciable differences

7. We have used the program mdraws written by Lorenzo Cappellari and Stephen P. Jenkins. See Cappellari and Jenkins (2006).

in the results, implying that the number of draws is large enough to provide an accurate approximation of the integral.

7.4.1 Estimation Results

Table 7.6 presents the estimation results for the equation for own work disability (equation [1]) and for DI receipt in the reference group (equation [7]). The estimates for the threshold equations (5) and (6) are given in table 7.7. Estimates for the vignette equations (equation [3]) are not of primary interest; they are presented and briefly discussed in table 7A.1 in appendix D.

Table 7.6 Estimation results for own work disability and receipt of DI benefits in the reference group

	Self-reported disability		Reference group disability	
	Coefficient	Standard error	Coefficient	Standard error
Age	–0.189	0.224	0.423	0.138
Age squared	0.017	0.020	–0.031	0.013
Medium education level	0.043	0.091	0.089	0.074
Higher education level	–0.085	0.105	–0.046	0.075
Female	0.003	0.075	–0.334	0.062
Age in reference group				
25–35	0.339	0.233	0.058	0.057
36–45	0.383	0.266	0.107	0.077
46–55	0.859	0.296	0.124	0.087
56–65	0.599	0.319	0.058	0.079
>65	0.636	0.333	–0.010	0.084
Medium education level in reference group	–0.246	0.093	0.016	0.027
High education level in reference group	–0.383	0.113	0.009	0.032
Northern provinces	0.061	0.124	–0.081	0.101
Eastern provinces	–0.026	0.104	–0.083	0.089
Western provinces	0.078	0.090	–0.285	0.073
Stroke	1.250	0.337	–0.029	0.244
Cancer	0.357	0.144	–0.193	0.157
Lung disease	0.661	0.142	0.281	0.132
Heart disease	0.825	0.132	–0.004	0.118
High blood pressure	0.029	0.086	0.069	0.075
Diabetes	0.408	0.180	0.118	0.154
Emotional problems	0.639	0.103	0.285	0.099
Arthritis	0.425	0.120	0.197	0.108
Problems with vision	0.076	0.178	0.035	0.163
Often pain	1.260	0.083	0.258	0.077
Intercept	–1.077	0.510	–1.378	0.333
ρ	0.053	0.040		
φ_{02}			1.338	0.051

Table 7.7 Estimation results of threshold equations

	Threshold shifts			
	γ_1		γ_2	
	Coefficient	Standard error	Coefficient	Standard error
Age	0.679	0.320	–0.104	0.037
Age squared	–0.055	0.027	0.009	0.003
Medium education level	0.157	0.102	–0.054	0.017
Higher education level	0.071	0.091	–0.040	0.016
Female	–0.373	0.187	0.024	0.016
Stroke	–0.110	0.320	–0.047	0.058
Cancer	–0.223	0.208	0.033	0.033
Lung disease	0.289	0.210	0.020	0.030
Heart disease	0.053	0.143	–0.065	0.029
High blood pressure	0.062	0.097	0.016	0.016
Diabetes	0.085	0.188	0.019	0.034
Emotional problems	0.238	0.186	–0.030	0.022
Arthritis	0.195	0.163	–0.021	0.022
Problems with vision	–0.038	0.198	0.031	0.036
Often pain	0.321	0.157	–0.025	0.017
Northern provinces	–0.128	0.128	0.022	0.021
Eastern provinces	–0.155	0.111	0.023	0.019
Western provinces	–0.347	0.167	0.046	0.017
Intercept	–2.068	0.898	0.259	0.108
μ	–0.968	0.023		
σ_ξ	0.733	0.072		
	Interactions			
	γ_1^R		γ_2^R	
	Coefficient	Standard error	Coefficient	Standard error
Age 35–64	0.034	0.042	–0.085	0.029
Age 65+	0.173	0.051	–0.072	0.028
Medium education level	0.049	0.045	–0.079	0.031
Higher education level	0.021	0.072	–0.009	0.040
Female	0.106	0.043	–0.111	0.031
Intercept	–1.333	0.554	0.249	0.059

Work Disability Self-Reports

The equation for own work disability in table 7.6 shows that there is virtually no gender difference (keeping other variables constant). Own work disability decreases with age until age fifty-six (age is measured in decades) and increases afterward; it is lower for higher-educated individuals than for respondents with low education. These effects are not statistically significant, however. Regional differences are not significant either. As expected, work limitations are significantly more frequent among individuals with serious health conditions, such as strokes, heart problems, cancer, diabetes, emotional problems, pain, and lung problems. Having a reference group with more medium or high education significantly reduces work disability.

DI Receipt in the Reference Group

The reference group DI receipt equation shows that the reported prevalence of DI receipt in the reference group increases with age until about retirement age (the estimated quadratic age function reaches a maximum at sixty-seven years of age). This is consistent with the fact that in the Netherlands, individuals over sixty-five typically do not receive DI benefits but receive state and occupational pensions instead. There is virtually no relation between DI receipt in the reference group and education. On the other hand, DI receipt in the reference group increases significantly with several health conditions (lung disease, emotional problems, pain), in line with the argument that people with a health problem will more often be acquainted with other people in poor health. Also in line with the raw data (table 7.4) is that females are significantly less likely to report to have DI-benefit recipients in their (female) reference group. Respondents in the western provinces of the country (the most urbanized region) are less likely to know people on disability benefits than respondents in the rest of the country.

The variables affecting the number of people on DI in the reference group are of interest in part because, as we shall see in the following, the number of people in the reference group significantly affects the thresholds used in evaluating work disability. For example, women know fewer people on DI and because of that will less easily say that a given health problem constitutes a work disability. Similarly, having pain increases the number of people on DI in one's reference group, and this makes people with pain "softer" in evaluating disability. These indirect effects come on top of the direct effects that gender and health conditions may have on the thresholds (see the following).

Thresholds

The results for the threshold equations are presented in table 7.7. We note that the overidentifying restrictions stemming from the fact that the reference group variables are not included in these equations do not get rejected ($\chi^2(12) = 11.382$; $p = .503$). The top panel presents estimates for the coefficients on individual characteristics in equations (5) and (6), while the bottom part shows the estimates of the coefficients of peer group DI receipt R_i^* interacted with education, age, and gender in both threshold equations. The estimates for the first threshold imply that women use lower thresholds than men with the same other characteristics and, thus, more easily regard a given health problem as work limiting. People with higher education are less likely to evaluate a given health problem as work limiting than low-educated respondents, but the educational differences are not statistically significant.

The age pattern is significant, and the age function has a maximum at about sixty-two years, implying that until age sixty-two, older people are "tougher," that is, less likely to call a condition work disabling. The only significant health condition is pain—respondents who often suffer from pain less easily evaluate a given health problem as a (mild or worse) work limitation, possibly because they are more used to performing work or daily activities in spite of the handicap of their health problem.

For the distance between the first and second threshold (γ_2), results are quite different. The age function has a minimum at fifty-six years of age (if $R_i^* = 0$), while higher education leads to a smaller distance between thresholds. Heart problems do the same; these are the only type of health problems with a significant effect. The estimates are difficult to interpret individually due to the complexity of the model, where the same variables appear in several equations.

The model parameters of primary interest are the coefficients γ_1^R and γ_2^R on peer group DI receipt R_i^*. Both have been specified as a function of education level, age, and gender (see the bottom panel of table 7.7). Consider first the estimated main effect and the interactions with education. For males under thirty-five with lower education, γ_1^R is estimated at –1.33; for otherwise identical individuals with medium education, the estimate is –1.28 (not significantly different from the –1.33 estimate), while for the higher educated, the estimate is –1.31. Females are significantly less influenced by DI receipt in their reference group than males; for example, for a lower-educated woman younger than thirty-five, the peer group effect is –1.23 (versus –1.33 for males). The significantly positive interaction of DI receipt in the reference group with the age dummy for sixty-five-plus shows that the response scale of individuals over sixty-five is less influenced by the number of DI recipients in their reference group than the response scale of younger individuals: the peer group effects are –1.16 for men over age sixty-five and –1.05 for women over age sixty-five.

Because the estimated value of γ_1^R is negative in all cases, the fraction of people who are on DI benefits in the reference group will unambiguously shift the reporting threshold for at least a mild working disability downward. In this sense, γ_1^R is the more critical parameter of the two. The estimates for γ_2^R show that the distance between the two thresholds increases with the number of friends and acquaintances on disability benefits, particularly for young males with the lowest education level. In simulations using the estimates of both γ_1^R and γ_2^R, we find that if the number of people on DI in the reference group increases, this raises both the fraction of those reporting they are somewhat limited and the fraction of those reporting they are moderately limited or worse, showing that the effect of R^* on γ_1^R dominates the effect on γ_2^R.

As mentioned earlier, we defined reference groups separately for men and women in the sense that for women we took the number of women on DI amongst female acquaintances and for men the number of male DI recipients among male acquaintances. One question is how sensitive the results in table 7.7 are to this particular specification of reference groups. To test this, we reestimated the model using a common definition of reference groups for both sexes.[8] The estimated effects of the number of people on DI in the reference group are even larger using the common reference by gender than with the benchmark definition used for table 7.7. A likelihood ratio test, however, indicates that the model with separate reference groups by gender for which we present the results is significantly better than the alternative model.

Covariance Structure of the Errors

Table 7.6 shows that the parameter ρ, the correlation between the error terms in the equations for own work disability (equation [1]) and DI receipt in the reference group (equation [7]) is small and insignificant. This is surprising because we would expect that work disability (and thus the unobserved factors driving it) positively affects the number of acquaintances on DI receipt.

Unobserved heterogeneity in thresholds is significant—the estimated standard deviation of ξ is 0.73 and is very accurately determined (σ_ξ in table 7.7). To judge its size, it can be compared to the amount of idiosyncratic noise in self-reports and vignette evaluations. The former has standard deviation 1 (by normalization), the latter has standard deviation 0.51 (see table 7A.1). Thus, unobserved heterogeneity in the thresholds explains about 35 percent of the unsystematic variation in self-reports and about 60 percent of the unsystematic variation in vignette evaluations.

The parameter μ is estimated at –0.97. Because $u_i^R = \omega_i^R - \mu\xi_i$ and $\text{Var}(u_i^R) = 1$ by means of normalization, we have $\text{Var}(\omega_i^R) = 0.50$. The implied correlation between ξ_i and u_i^R is equal to 0.71. The sign of μ implies that respondents who use relatively high thresholds for answering questions about their own work limitations (given their observed characteristics) will tend to use relatively low thresholds when asked for DI prevalence in the reference group. Thus, someone who is unlikely to refer to a health problem as work limiting has a tendency to consider work limitations as more of an exception and will sooner consider a given number of people on DI in the reference group as "many."

8. All respondents were asked both the number of men and the number of women on DI in their reference group. To form a common definition for men and women, we used the maximum of the two. Thus, if for an individual respondent there were a lot of individuals of one gender who were more than somewhat limited, that is the value that applies.

Table 7.8 Model predictions of self-reported work disability and reference group disability

	Disability in the reference group (%)			
Self-reported work disability	None	Very few	A few/many	Total
Not limited	61.3	32.7	6.0	100.0
	74.0	63.2	53.9	68.6
Mildly limited	50.6	39.7	9.7	100.0
	16.6	20.9	23.8	18.7
Moderately, severely, and extremely limited	42.1	44.6	13.3	100.0
	9.4	15.9	22.3	12.7
Total	57.1	35.2	7.7	100.0
	100.0	100.0	100.0	100.0

Notes: Data are weighted. N = 1,764 (estimation sample).

7.4.2 Model Performance

Table 7.8 provides a simple way of checking the fit of the model. Its structure is similar to that of table 7.5, but it reports simulated frequencies using the model instead of actual frequencies in the data. Comparing table 7.8 with table 7.5 suggests that the fit of the model is fairly good; judging from the marginal distributions, the model does a good job in replicating reported reference group DI receipt; it does a slightly worse job in reproducing the distribution of self-reported disability. The biggest deviation between the data and the model predictions occurs in the middle category (mildly limited). According to the data, 22.4 percent of the respondents classify themselves as mildly limited (table 7.5), whereas the model predicts 18.7 percent in that category (table 7.8).

7.4.3 Simulation of Reference Group Effects

One way to gauge the strength of the reference group effects is to artificially vary the number of people on DI in an individual's reference group and then to evaluate how this affects the prevalence of self-reported work limitations. We do this by varying the intercept in the equation for the number of people in the reference group on DI (equation [7]) and then simulate the reports of DI-benefit receipt in the reference group and the prevalence of self-reported work disability induced by that new level of reference group DI receipt.

Figure 7.1 shows the results for both the full sample and for the sample broken down by education. In each picture, the horizontal axis is the percentage of respondents who say that they know at least a few DI-benefit recipients, with the vertical lines representing the sample (or subsample) percentages (except the left vertical line in the first figure, see the following). The vertical axis represents the percentage who report that they suffer from

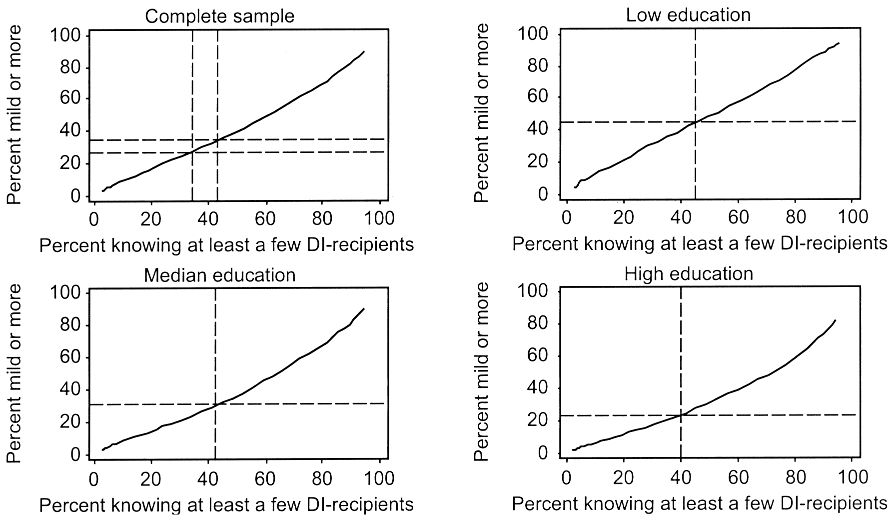

Fig. 7.1 Self-reported work disability and reference group DI

at least a mild work limitation; the horizontal line indicates the (sub)sample percentage (except the lower line in the first figure).

The graphs in the figures illustrate the sensitivity of reporting a work disability to DI receipt in the reference group. In line with the estimation results in table 7.6, the level of the curve is highest for the low educated and lowest for those with a high education level. This difference in levels implies that at the same level of perceived reference group DI benefit receipt, lower-educated respondents are more likely to report at least a mild work limitation than respondents with middle or higher education. In all cases, there is a notable peer group effect of DI receipt in the reference group on the probability to report a work disability: if the respondent knows more people on DI benefits, the chances of reporting a disability increase substantially.

To illustrate the size of the effect, in the picture for the full sample, additional horizontal and vertical lines have been drawn, both below the sample averages. The horizontal line is based on the finding of Kapteyn, Smith, and van Soest (2007) that if US scales are assigned to Dutch respondents, self-reported work limitations in the Netherlands would fall by 21 percent.[9] This second horizontal line can thus be interpreted as self-reported work limitations in the Netherlands if the Dutch respondents with the Dutch work limitations would use the American response scales. The second vertical line shows that if the percentage of individuals saying they know at least a few DI-benefit recipients in their reference group were to move from its

9. This is the finding in their benchmark model; the percentage varies somewhat depending on which model specification is chosen.

simulated sample mean of 42.9 percent to about 33.9 percent (the left most vertical line), this would move the scales used by the respondents enough to reach the US scales.

7.5 Concluding Remarks

Most people do not live in social isolation. Instead, they interact repeatedly with family, friends, and neighbors. As a consequence of those pervasive interactions, they allow themselves to be transformed in many ways, a transformation of which they may often be unaware. One type of transformation involves the formation of social norms about what normal or acceptable behavior might be. These social norms then fix the scales that they may be using in responding to questions about their own behaviors and current situations. If they had different neighbors and friends, their self-descriptions about their lives may well be quite different. While this may be true within a country where there exists a shared history and culture, it is especially likely to be the case when cross-national comparisons are made.

In this chapter, we test the importance of these types of social interactions using a specific application—the probability that people self-label themselves as work disabled. We estimated a model of self-reported disability with an emphasis on how the reporting of disability is affected by the prevalence of DI receipt in one's reference group. We find an effect in the hypothesized direction—larger reported numbers of people in one's reference group on DI increase the likelihood of seeing oneself as having a work disability.

An alternative interpretation of our main finding has been suggested. It is reasonable to expect that people with lenient standards also perceive more of their friends as disabled. This could invalidate our interpretation if our reference group variable was the number of friends perceived as disabled, but this is not the case—it refers to the number of friends receiving disability insurance benefits—an objective criterion.

These findings are suggestive of how policy programs affect social norms. If a policy makes receipt of DI benefits more attractive or easier (e.g., by loosening eligibility requirements) thus increasing the number of DI recipients, this changes social norms. Individuals are now more likely to label a given health condition as work limiting and the prevalence of self-reported work will rise.

There are of course alternative reasons why self-reported disability and reported DI-benefit receipt in one's reference group would be correlated. Our model is designed to capture many of these reasons. These include the possibility that individuals with a work disability are more likely to associate with others who suffer a similar fate. First, we allow for a considerable number of observable covariates in common, which by itself will generate correlation between self-reported disability and reported DI-benefit receipt in one's reference group. But we also allow for correlation between the errors in the reference group equation and the equation predicting the probability that someone is work disabled.

Even within this reasonably general model, we find a direct effect of the number of people in one's reference group on disability programs on the probability one considers oneself work disabled. The effects that we estimate are sufficiently strong to explain a good deal of the higher rates of self-reported work disability in the Netherlands compared to the United States. The Dutch population appears to have much more lenient thresholds about what constitutes a work disability (Kapteyn, Smith, and van Soest 2007). The results in this chapter suggest that this tendency stems from the fact that the Dutch are much more likely to know people on work disability programs, a direct consequence of the far more generous programs in the Netherlands as well as its more lenient rules for program eligibility at the time of the survey.

Appendix A
Vignette Questions

Vignettes for Affect

1. [Henriette] 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.

2. [Jim] enjoys work very much. He feels that he is doing a very good job and is optimistic about the future.

3. [Tamara] 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.

4. [Eva] 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.

5. [Roberta] feels depressed most of the time. She weeps frequently at work and feels hopeless about the future. She feels that she has become a burden to her coworkers and that she would be better dead.

Vignettes for Pain

1. [Katie] occasionally feels back pain at work, but this has not happened for the last several months now. If she feels back pain, it typically lasts only for a few days.

2. [Catherine] 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. [Yvonne] has almost constant pain in her back, and this sometimes prevents her from doing her work.

4. [Jim] has back pain that makes changes in body position while he is working very uncomfortable. He is unable to stand or sit for more than half an hour. Medicines decrease the pain a little, but it is there all the time and interferes with his ability to carry out even day-to-day tasks at work.

5. [Mark] 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 and when holding and lifting things at work.

Vignettes for CVD

1. [Trish] is very active and fit. She takes aerobic classes three times a week. Her job is not physically demanding, but sometimes a little stressful.

2. [Norbert] 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.

3. [Paul]'s family has a history of heart problems. His father died of a heart attack when Paul was still very young. The doctors have told Paul that he is at severe risk of having a serious heart attack himself and that he should avoid strenuous physical activity or stress. His work is sedentary, but he frequently has to meet strict deadlines, which adds considerable pressure to his job. He sometimes feels severe pain in chest and arms and suffers from dizziness, fainting, sweating, nausea, or shortness of breath.

4. [Tom] 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. His job is not physically demanding, but sometimes it can be hectic. He does not get along with his boss very well.

5. [Dan] has undergone triple bypass heart surgery. He is a heavy smoker and still experiences severe chest pain sometimes. His job does not involve heavy physical demands, but sometimes at work he experiences dizzy spells and chest pain.

Appendix B

Reference Group Questions

The questions are preceded by the following introduction: *The following questions concern your circle of acquaintances, that is, the people with whom you associate frequently, such as friends, neighbors, acquaintances, or maybe people at work.*

- If you think of your circle of acquaintances, into which age category do MOST of these people go? Please select the answer that is closest to reality.
  Age (in years) is mostly: 1. under 16; 2. 16–20; 3. 21–25; 4. 26–30; 5. 31–35; 6. 36–40; 7. 41–45; 8. 46–50; 9. 51–55; 10. 56–60; 11. 61–65; 12. 66–70; 13. 71 or over
- Which level of education do most of your acquaintances have?
  1. primary education; 2. junior vocational training; 3. lower secondary education; 4. secondary education/preuniversity education; 5. senior vocational training; 6. vocational colleges/first-year university education; 7. university education
- If you think of the men among your acquaintances, how many of them are on DI?
  1. Nobody; 2. Very few; 3. A few; 4. Many
- If you think of the women among your acquaintances, how many of them are on DI?
  1. Nobody; 2. Very few; 3. A few; 4. Many

Appendix C

Likelihood Contributions

Compared to the models in King et al. (2004) and Kapteyn et al. (2007), there are two complications: the thresholds now depend on an unobserved variable R^* and upon an unobserved heterogeneity term ξ. Replacing R^* using equation (7) and exploiting equations (5) and (6) gives:

(C1) $$\tau_1 = V\gamma_1 + \gamma_1^R X^R \beta^R + \xi + \gamma_1^R(u^R + \mu\xi),$$

(C2) $$\tau_2 = \tau_1 + e^{V\gamma_2 + \gamma_2^R X^R \beta^R + \gamma_2^R(u^R + \mu\xi)}.$$

Equations (1) and (2) imply

(C3) $$Y = j \text{ if } \tau_{j-1} - X\beta < \varepsilon < \tau_j - X\beta.$$

Similarly, for the vignette evaluations, we get:

(C4) $$Y^l = j \text{ if } \tau_{j-1} - \theta^l - \delta F^l < \varepsilon^l < \tau_j - \theta^l - \delta F^l.$$

The probability of observing a certain reference group category follows from equation (9):

(C5) $$R = j \text{ if } \phi_{0,j-1} - X^R\beta^R < u^R < \phi_{0j} - X^R\beta^R$$

Let the reported reference group variable be r, the observed work disability self-report y, and the observed vignette evaluations $y^1, \ldots, y^L$. Then the likelihood contribution of a given respondent can be written as a two-dimensional integral over the values of u^R that result in $R = r$ and all possible values of ξ:

$$\text{(C6)} \quad \int_{-\infty}^{\infty} \int_{\phi_{0,j-1} - X^R\beta^R}^{\phi_{0j} - X^R\beta^R} P(Y = y \mid u^R, \xi) \prod_{l=1}^{L} P(Y^l = y^l \mid u^R, \xi) f(u^R \mid \xi) du^R \frac{1}{1\sigma_\xi} \phi\left(\frac{\xi}{\sigma_\xi}\right) d\xi,$$

where ϕ is the standard normal density, and f is the conditional density of u^R given ξ, which is univariate normal. Of course, the crucial point here is that, conditional on u^R and η, all vignette evaluations and the self-report are mutually independent, allowing for the factorization in equation (C6). The conditional probabilities in equation (C6) follow from equations (C3) and (C4), together with the normality assumptions on the error terms, implying that the ε^l are independent of ε, ξ, and u^R but that $\varepsilon|(u^R, \xi) \sim N(\rho u^R, 1 - \rho^2)$:

$$P(Y = y \mid u^R, \xi) = \Phi\left(\frac{\tau_y - X\beta - \rho u^R}{\sqrt{1-\rho^2}}\right) - \Phi\left(\frac{\tau_{y-1} - X\beta - \rho u^R}{\sqrt{1-\rho^2}}\right)$$

$$P(Y^l = y^l \mid u^R, \xi) = \Phi\left(\frac{\tau_{y^l} - \theta^l - \delta F^l}{\sigma_\upsilon}\right) - \Phi\left(\frac{\tau_{y^l-1} - \theta^l - \delta F^l}{\sigma_\upsilon}\right),$$

where the $\tau_{...}$ are given by equations (C1) and (C2) and depend on ξ and u^R.

Appendix D
Estimates of the Vignette Equation (Equation [3])

The dummy coefficients in table 7A.1 reflect the average severity of the work limitations described in the vignettes. One can relate the dummy coefficients θ^l, $l = 1, \ldots, 15$ to the relative frequencies in table 7.1—vignettes that are evaluated as more severely on average have higher coefficients. The estimate of δ, the coefficient of the dummy for a female vignette name, is small and insignificant. The estimated idiosyncratic variation in vignette evaluations σ_υ (independent across vignettes) is smaller than the unsystematic variation in self-assessments ($\sigma_\varepsilon = 1$, by means of normalization). Still, the idiosyncratic terms ε_υ are large enough in comparison to the differences in the estimated coefficients on the vignette dummies θ^l to explain that the same vignettes are often ranked in different ways by different respondents—in line with what is seen in the data.

Table 7A.1 Estimates of vignette equations

Vignette	Coef.	s.e.
1	—	
2	–1.295	0.067
3	0.716	0.041
4	0.541	0.032
5	1.246	0.066
6	0.302	0.025
7	0.758	0.039
8	1.589	0.079
9	1.549	0.077
10	1.360	0.069
11	–0.962	0.050
12	0.685	0.037
13	1.149	0.060
14	0.487	0.030
15	0.795	0.044
σ_υ	0.515	0.024
δ	–0.006	0.019

References

Aarts, L., R. Burkhauser, and P. de Jong. 1996. *Curing the Dutch Disease: An International Perspective on Disability Policy Reform.* Advanced Textbooks in Economics. Avebury, UK: Aldershot.

Alessie, R., and A. Kapteyn. 1991. "Habit Formation, Interdependent Preferences and Demographic Effects in the Almost Ideal Demand System." *Economic Journal* 101:404–19.

Angelucci, M., and G. D. Giorgi. 2009. "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review* 99 (1): 468–508.

Audretsch, D., and M. Feldman. 1996. "R&D Spillovers and the Geography of Innovation and Production." *American Economic Review* 86:630–40.

Bala, V., and S. Goyal. 1998. "Learning from Neighbours." *Review of Economic Studies* 65:595–622.

Banerjee, A. V. 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics* 107:797–817.

Banks, J., A. Kapteyn, J. P. Smith, and A. van Soest. 2008. "Work Disability is a Pain in the *****, Especially in England, the Netherlands, and the United States." In *Health at Older Ages: The Causes and Consequences of Declining Disability Among the Elderly,* edited by D. Cutler and D. Wise, 251–94. Chicago: University of Chicago Press.

Becker, G. S. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 83 (6): 1063–93.

Bound, J. 1991. "Self-Reported versus Objective Measures of Health in Retirement Models." *Journal of Human Resources* 26 (1): 106–38.

Bound, J., and R. Burkhauser. 1999. “Economic Analysis of Transfer Programs Targeted on People with Disabilities.” In *Handbook of Labor Economics.* Vol. 3C, edited by O. Ashenfelter and D. Card, 3417–3528. Amsterdam: Elsevier.

Cappellari, L., and S. Jenkins. 2006. “Calculation of Multivariate Normal Probabilities by Simulation, with Applications to Maximum Simulated Likelihood Estimation.” IZA Discussion Paper no. 2112. Bonn, Germany: Institute for the Study of Labor.

Case, A., and L. Katz. 1991. “The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth.” NBER Working Paper no. 3705. Cambridge, MA: National Bureau of Economic Research.

Duesenberry, J. 1949. *Income, Saving and the Theory of Consumer Behavior.* Cambridge, MA: Harvard University Press.

Duflo, E., and E. Saez. 2003. “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment.” *Quarterly Journal of Economics* 118 (3): 815–42.

Ginther, D., R. Haveman, and B. Wolfe. 2000. “Neighborhood Attributes as Determinants of Children’s Outcomes: How Robust are the Relationships?” *Journal of Human Resources* 35 (4): 603–42.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman. 1996. “Crime and Social Interactions.” *Quarterly Journal of Economics* 111 (2): 507–48.

———. 2000. “The Social Multiplier.” *Journal of the European Economic Association* 1 (2–3): 345–53.

Hanushek, E., J. Kain, J. Markman, and S. Rivkin. 2003. “Does Peer Ability Affect Student Achievement?” *Journal of Applied Econometrics* 18 (5): 527–44.

Haveman, R., and B. Wolfe. 2000. “The Economics of Disability and Disability Policy.” In *Handbook of Health Economics.* Vol. 1B, edited by J. Newhouse and A. Culyer, 995–1051. Amsterdam: Elsevier.

Kapteyn, A., J. P. Smith, and A. H. van Soest. 2007. “Vignettes and Self-Reports of Work Disability in the U.S. and the Netherlands.” *American Economic Review* 97 (1): 461–73.

Kapteyn, A., S. van de Geer, H. van de Stadt, and T. Wansbeek. 1997. “Interdependent Preferences: An Econometric Analysis.” *Journal of Applied Econometrics* 12 (6): 665–86.

King, G., C. Murray, J. Salomon, and A. Tandon. 2004. “Enhancing the Validity and Crosscultural Comparability of Measurement in Survey Research.” *American Political Science Review* 98 (1): 567–83.

Kuhn, P., P. Kooreman, A. R. Soetevent, and A. Kapteyn. 2011. “The Effects of Lottery Prizes on Winners and Their Neighbors. Evidence from the Dutch Postcode Lottery.” *American Economic Review* 101 (5): 2226–47.

Manski, C. F. 1993. “Identification of Endogenous Social Effects.” *Review of Economic Studies* 60 (3): 531–42.

Topa, G. 2001. “Social Interactions, Local Spillovers and Unemployment.” *Review of Economic Studies* 68 (2): 261–95.

Veblen, T. 1899. *The Theory of the Leisure Class.* New York: Dover.

Woittiez, I., and A. Kapteyn. 1998. “Social Interactions and Habit Formation in a Model of Female Labour Supply.” *Journal of Public Economics* 70:185–205.