Disability Insurance and Health in Europe and the US

Enrica Croda
Department of Economics
Ca' Foscari University of Venice
Cannaregio, 873 - 30121 Venice (Italy)
enrica.croda@unive.it

Jonathan Skinner
Department of Economics
Dartmouth College HB 6106
Hanover NH 03755
jon.skinner@dartmouth.edu

July 2010 – ***Very Preliminary***

Abstract

Rising fiscal pressure from expanding disability insurance (DI) programs has increased pressure on governments to scale back benefits, but there is concern that such reductions will leave workers in poor health outside the safety net. We first develop a model that captures this tension between budgetary costs and the health of DI applicants, but also allows for cross-country differences in two key parameters: the importance of employment opportunities in determining eligibility, and the efficiency in screening for health-related disabilities. We test the model using 2004 data on individuals from 12 countries in the Survey of Health, Ageing, and Retirement in Europe (SHARE) and the Health and Retirement Study (HRS). First, we confirm that there is large variation across countries in the share of those aged 50-64 enrolled in disability insurance programs, ranging from 2.3 percent in Greece and 3.9 percent in France to more than 14 percent in Sweden, Denmark, and the Netherlands. Second, the evidence does not favor the conventional "tradeoff" model: DI enrollees in low-enrollment countries are no sicker or depressed than DI enrollees in high-enrollment countries. Finally, this puzzling pattern can be explained in part by differences across countries in the weight they place on labor market opportunities, and how well they screen for illness. The predictive ability of the model is weak, but it does suggest that policy reforms should look beyond simply loosening or tightening eligibility requirements to helping workers experiencing chronic disability, pain, and poor employment prospects remain in the labor force.

We are grateful to financial support from NETSPAR and the National Institute on Aging (PO1 AG019783), and thank without implicating Lans Bovenberg, Dimitris Christelis and Douglas Staiger for helpful comments. This paper uses data from Release 2.3.0 of SHARE 2004. SHARE data collection in 2004-2007 was primarily funded by the European Commission through its 5th and 6th framework programmes (project numbers QLK6-CT-2001- 00360; RII-CT- 2006-062193; CIT5-CT-2005-028857). Additional funding by the US National Institute on Aging (grant numbers U01 AG09740-13S2; P01 AG005842; P01 AG08291; P30 AG12815; Y1-AG-4553-01; OGHA 04-064; R21 AG025169) as well as by various national sources is gratefully acknowledged (see http://www.share-project.org for a full list of funding institutions).

1. Introduction

There are large variations across countries in the percentage of GDP devoted to disability payments, ranging in western countries from 0.4 percent in Canada to 2.5 percent in Sweden (OECD, 2009). Most countries are critically concerned about rising rates of disability insurance enrollment, and the resulting pressure on public sector budgets in European countries and the United States (European Commission, 2006; OECD, 2003; McVicar, 2008; Autor and Duggan, 2006).

In response, some governments have restricted eligibility and reduced payment rates (Euwals, et al., 2009), but these reforms in turn have caused concerns about restricted access to DI for people with very serious disabilities that prevent them from working. The U.S. disability insurance program is particularly notable for the very long waiting period and extensive appeals for people with what appear to be serious disabilities (Eckholm, 2007). By contrast, the Swedish system allows for disability to be defined in percentage terms and provides smaller awards for those less disabled. Thus governments adopt very different approaches to navigating between Type I error (rejecting a truly disabled worker) and Type II error (allowing non-disabled worker to receive DI). In this view, the size and growth of DI programs should reflect where countries choose to draw the line in determining whether specific workers are sufficiently disabled to become eligible for DI.[1]

We build on this intuition by creating a simple model of DI application and enrollment that allows for systematic variation across countries in their objective function – how much to insure against poor health or poor employment opportunities – and in the apparent randomness of

[1] Some countries have tried to sidestep this question by defining disability in percentage terms and awarding smaller awards for less severe disabilities, but this approach has also expanded the potential pool of enrollees at the same time.

the application process. The first implication of the model is that, as above, countries setting stricter eligibility for health-related disability will experience both lower DI enrollment rates, and conditional on receiving DI, enrollees will also be substantially sicker than average (and much sicker than those not on DI). The second and less intuitive implication is that difference in how countries make tradeoffs between health and employment blurs and could even offset the first implication of the negative association between the generosity of the program and the health of its enrollees. Finally, country-level systems with more "error" in terms of DI acceptance or denial based on idiosyncratic characteristics of applicants will further attenuate health-related differences between the DI and the non-DI population.[2]

We use data from the Survey of Health, Ageing, and Retirement in Europe (SHARE) and the Health and Retirement Study (HRS) to compare patterns of health and DI participation for people aged 50-64 across eleven countries in Europe and the U.S. We first demonstrate the *lack* of correlation between the size of the DI program and the average or relative self-reported health or depression scores of those in the DI program. In other words, the first implication of the simple model, that country DI programs are similar except for where they draw the line on who is sufficiently sick for eligibility, is not supported by the data.

We then use the micro-level SHARE and HRS data to shed light on differences across countries in characteristics of DI enrollees. In a previous study, Borsch-Supan (2007) has shown that differences across countries in their DI enrollment rates are driven by institutional differences, and not by sociodemographic differences across countries. European countries appear to place a greater weight on the lack of market opportunities, as proxied by education, while in the US there is no impact of market opportunities on DI enrollment once one controls

[2] The key assumption we make is that self-reported health or depression is a good measure of health and ability to work. We discuss this limitation below.

for self-reported health. As well, the magnitude of the coefficients predicting DI enrollment are supportive of the view that countries also differ with regard to the variance in the selection process. While Denmark and Sweden place a similar (high) relative weight on insuring against a lack of employment opportunities, the Danish system is far better at screening out people without debilitating illnesses by keeping them in the labor force, a feature noted independently by a recent OECD (2009) study.

What's going on in the U.S.? It has among the strictest *de facto* eligibility rules and require strong documentation of a medical illness (rather than tying benefits to workplace productivity, as in many European countries), while exhibiting very high rejection rates (OECD, 2009). And people in the U.S. program systematically report worse health status and a greater level of depression relative to the non-disabled population. Yet the enrollment rate is not much different from median of the 11 European countries. In part, Americans may be somewhat sicker (e.g., Banks et al., 2006) but it may also reflect an endogenous process whereby individuals need to signal severe disability in order to qualify for DI (Atlas and Skinner, 2009).

The policy implications of these results are both discouraging and encouraging. Discouraging because the association between DI enrollment and health status appears to do such a poor job of targeting people in the poorest health, a result also found in a longitudinal setting by Borsch-Supan (2008). More encouraging, however, is that countries do not appear to be facing the rigid tradeoff implied in the simple model of Type 1 versus Type 2 error in determining eligibility for DI. Instead, the wide variability across countries suggests a much larger scope for improving the efficiency of DI programs by intervening quickly and providing supportive employment to avoid the worst-case outcome of a permanent transition to long-term chronic pain and disability (Burns, 2007; Drake et al., 2009).

2. The Model

We begin with a model of disability where the decision to go on disability is jointly determined by whether individual i in country j chooses to apply for benefits (A_{ij}) and whether the applicant is approved by the country DI program (Y_{ij}). Consider first the individual's decision to apply:

$$
\begin{aligned}
& A_{ij}^{*} = \Omega_{j} Z_{ij} + \phi_{i} h_{ij} + \pi_{j} w_{ij} + \varepsilon_{i} \\
(1) \qquad & A_{ij} = 1 \quad if \quad A_{ij}^{*} > 0 \\
& A_{ij} = 0 \quad otherwise
\end{aligned}
$$

where A_i* is a linear index which in turn depends on exogenous factors Z_i, health status h_i, and market wage opportunities w_i, while A_i is one if individual i applies to receive DI benefits, and zero otherwise.

The individual only becomes eligible for DI only if the DI authority in country j approves the application, where again we assume that there is an application index Y_{ij}*

$$
\begin{aligned}
& Y_{ij}^{*} = \beta_{j} X_{ij} + \alpha_{j} (h_{ij} - H_{j}) + \gamma_{j} (w_{ij} - W_{j}) + u_{ij} \\
(2) \qquad & Y_{ij} = 1 \quad if \quad A_{ij}^{*} > 0 \\
& Y_{ij} = 0 \quad otherwise
\end{aligned}
$$

In this specification, the disability board observes a subset X of the total exogenous individual level variables Z, for example age or gender. Countries also differ with regard to their overall generosity, which is a combination of the "reference" levels of health H_j and potential labor earnings W_j as well as the coefficients reflecting the relative importance of health (α) and labor market opportunities (γ) in predicting whether the DI program would approve the application. Note that we can further summarize the country-specific component as

$$
\begin{aligned}
& V_{ij}^{*} = \beta_j X_{ij} + \alpha_j h_{ij} + \gamma_j w_{ij} + u_{ij} \\
(2') \qquad & Y_{ij} = 1 \quad if \;\; V_{ij}^{*} > C_j = \alpha_j H_j + \gamma_j W_j \\
& Y_{ij} = 0 \;\; otherwise
\end{aligned}
$$

That is, country-level policies can be summarized as a function of the relative weight for employment opportunities and health, as well as the overall level of generosity C_j.[3] Note also that countries may differ with regard to the variance of the error term, where $\sigma_j^2 = \mathrm{Var}(u_{ij})$, so that (for example) selection processes with considerable randomness exhibits a much larger variance. Finally, disability enrollment D_{ij} is the product $A_{ij}\, Y_{ij}$, where D_{ij} is equal to one if the applicant is approved.

There are several terms in this equation that may be normalized; probit models by definition impose unit variance. Thus the model is informative about the "true" coefficients divided by the standard deviation of σ_j, allowing us to make inferences about the relative weights for health versus employment opportunities (that is, the ratio or the log difference of the coefficients) and the degree to which selection is based on self-reported health or depression. In the extreme where the selection process is nearly random, the magnitude of both coefficients would be small, meaning that neither labor market opportunities nor health were predictive of DI enrollment; equivalently, we might expect the average health or employment prospects of DI enrollees to be similar to those not on DI.

In theory, this model can be estimated using a partial-observation probit model (Poirier, 1980) where one only observes a positive value for both the application being filed and the application being approved by the DI program. However, Meng and Schmidt (1985) have pointed out that the efficiency of the estimator is poor, and at this stage the poor identification of the different equations makes such an approach problematic. (The next version of this paper will

[3] One needs to normalize elements of Ω in order to uniquely identify C.

be able to estimate each equation separately because of newly released SHARELIFE/SHARE Wave 3 data on whether the respondent ever applied for DI.) We instead assume a single-equation "reduced form" estimation model.[4] By implicitly assuming that individuals who apply for DI choose using the same parameters in the Y equation by the DI program to determine eligibility, we will (for now) interpret the reduced form coefficients as reflecting the institutional parameters of the DI program.

Figure 1 illustrates the simple model for the health and wage (or market opportunity) dimensions. The ellipse drawn in Figure 1 represents the distribution of health (h) and wage opportunities (w) in this general population for Country Z (conditional on X). We draw an ellipse rather than a circle to reflect the observed correlation between health and wages.

Consider first the red line (mm') with an intercept on the Y axis equal to $[C_j - X\beta]/\gamma$, and with a slope equal to $-\alpha/\gamma$. (Since we would expect both α and γ to be negative, this slope should also be negative.) In this deterministic model, everyone below the line mm' should be on DI insurance, while everyone above the line should not be on DI. One implication of this model also is that, conditional on being in DI, those with worse market opportunities should also be in better health. The point A represents the mean value of both wages (its height along the Y axis) and health (its distance along the X axis) for people who are enrolled in DI, while B shows the corresponding means for those who are not on DI.

Consider now a different country with more liberal rules for DI enrollment, reflected in a higher C_j (and thus easier to qualify for given that α and γ are both negative), shown in the blue line (nn'). This new policy will result in a larger set of people eligible for DI, with a resulting rise in average values both for health status and wages conditional on eligibility, as shown by the

[4] One can show that such a "reduced form" is a linear combination of the application coefficients (for A) and the approval coefficients (for Y) in a logit model (the analytics are more difficult in the probit model).

upward (and outward) shift in the conditional mean from A to A'. Thus in this simple model, one might expect that more generous DI plans would lead to more people becoming eligible (as in Börsch-Supan, 2007) which in turn leads to a generally less severely disabled group of DI recipients with better market opportunities.

If countries differ with respect to underlying health, this simple implication of the model might not hold, as shown in Figure 2 which allows for countries to exhibit different levels of self-reported or documented health levels (e.g., Banks, et al., 2006). If country R has a more restrictive DI policy (mm') than country P (nn'), but country P is less healthy, corresponding to the shifted distribution shown in Figure 2, then the conditional mean health for country P (A') for DI participants could still be below the conditional health for country R (A), despite the fact that a larger fraction of the ellipse (and hence the distribution of workers) are enrolled in DI in country R.

One approach to sidestep this potential confounder is to compare the *difference* in self-reported health between the disabled and non-disabled groups, thereby implicitly controlling for country-level differences in the mean values of health measures. As well, this also controls for the findings that people in different countries seem to have different norms about what constitutes poor health (Kapteyn, et al., 2007). For example, in Figure 2 a movement to a more generous DI program results in a shrinkage in the difference between the means (as measured in terms of health on the horizontal axis) in Country P (that is, A' versus B') in comparison to the wider spread in health for the DI versus the non-DI groups in Country R (A versus B).

This property certainly holds in the graph as drawn, but it is not always true, as it depends on the shape of the density function for health and the location of the cut-point. (One simple counter-example – if the cut-point of a normal distribution went from the 90^{th} to the 95^{th}

percentile of the distribution of health, the mean health for the vast majority of people receiving benefits would respond more sluggishly than the mean health for the shrinking group of non-recipients.) However, simulations using the normal distribution suggested that this is true empirically, and in general one would expect to find this weaker condition – that the difference in health status shrinks as more people become eligible – to hold when enrollment rates are 30% or less. Note also that this approach can potentially control for differences in the structure of the questionnaire between SHARE and HRS, as we shall see in the empirical section.

Another possible explanation for our empirical puzzle is that countries differ systematically with regard to their relative weights placed on health and potential wage rates or market opportunities in judging eligibility for DI. Figure 3 shows a scenario in which the wage and health distribution is the same for two countries, but they differ with regard to the relative emphasis placed on health versus market opportunities (ie the ratio of the coefficients α/γ vary). In this case, the more steeply sloped line nn', which compared to mm' places greater emphasis on health qualifications (that is, the "medical model" of disability), could result in a sicker group of people (both in a relative and absolute sense) but with a higher fraction who are actually enrolled in DI.[5]

A final possibility relates to the size of the error term in the decision, as noted above. To this point, the graphs have delineated the population into the two groups without any uncertainty or randomness. In the presence of randomness in the selection process, there would be more crossover, which would of course attenuate the health differences between the two groups and reduce the predictive power of health and labor market opportunities in determining whether the individual enrolled in DI.

[5] See OECD (2003) for a discussion of disability models.

3. The Data

We use the SHARE and HRS data on people age 50-64. Properties of the SHARE data, such as response rates and sample sizes, have been reported elsewhere (e.g., Börsch-Supan, 2007). The HRS is a similar longitudinal study, although for this analysis (and to allow comparability with the SHARE data) we consider just the cross-sectional data from 2004.

About two-thirds of the variables in SHARE are identical to variables in HRS, and most of the remainder is fairly comparable (Börsch-Supan, 2007). Some transformations of the original variables have been necessary to ensure close comparability between the two data sets. Self-reported health (excellent, very good, good, fair, poor) lined up well, although of course people in different countries may interpret the categories differently (e.g., Kapteyn, et al., 2007).

The construction of a closely comparable measure of depression required some variables transformations. First, both the SHARE and the HRS asked only a subset of the original CES-D standard depression items. The SHARE survey (in the "drop-off" to the main CAPI questionnaire) asked 14 of the original CES-D items, while the HRS asked only 9 of them. We matched the 9 questions present in both data sets. More importantly, the format of the SHARE and HRS answers to the CES-D questions is different. The HRS used a simple "yes-no" response format, while SHARE used a four-level frequency response format, proposing four different categories: "almost all of the time", "most of the time", "some of the time", "almost none of the time." [6] We assigned the first two to "yes" and the second two to "no." Our modified score ranged from -3 to 6, with -3 corresponding to "best case" mental states and 6 the worst. However, this mapping is not perfect, as shown by Steffick (2000) who compared the two-answer and four-answer combinations during the year in the HRS when both sets of responses

[6] HRS in the very first wave used a four-level frequency response (but slightly different phrasing than SHARE: rarely/none of the time, some of the time, most of the time, all/almost all of the time) but then switched to the yes/no format in order to simplify telephone administration.

were provided, since relatively few of the "some of the time" group answered "yes" when given the chance.[7] We adjust for this problem by considering primarily the difference in our adjusted depression score between DI enrollees and non-DI enrollees. This adjustment controls for differences in the mean value, but may still be imperfect when the distributions differ.

Education was split into primary, secondary, and tertiary. For SHARE, we relied on the ISCED-97 coding provided in the generated variables file.[8] For the HRS, primary corresponded to 11 years of education or less, secondary was 12 years, and tertiary was more than 12 years of education.

4. Results

In a seminal study of disability using the SHARE and HRS data along with the ELSA data from the UK, Börsch-Supan (2007) noted the wide variation across countries in enrollment rates for the 50-65 population, and suggested that most of the variation could be attributed to institutional country-level differences in eligibility and compensation levels.

How should these variations affect the average level of disability across these countries? The simplest view of a disability program is one which, in an ideal world, leads to the most severely disabled receiving benefits first, and then as eligibility expands, the program moves

[7] In an experimental module of HRS wave 2, a subset of respondents was asked both the four-level frequency questions used in HRS wave 1 and the then new yes/no response scale. Steffick (2000) finds some disagreement between the two forms of the scale and shows that the major sources of discrepancy from collapsing the four-level response categories into yes/no responses are the respondents that report "some of the time" on the frequency response. It follows that designating "some of the time" respondents as "yes" overstates the endorsement of the item, while designating them as "no" understates the prevalence.

[8] The SHARE generated variables file provides the 1997 International Standard Classification of Education (ISCED-97) coding. We combined the ISCED-97 codes 0 (none), 1 (primary education), 2 (lower secondary education), into one category ("primary"), the codes 3 (upper secondary education) and 4 (post-secondary, non-tertiary education) into another category ("secondary"), and categories 5 (first stage of tertiary education) and 6 (second stage of tertiary education) into yet another category ("tertiary"). See separate Data Appendix for further details.

down the severity curve so that successively less disabled people becomes eligible. (While we recognize the distinction between health and disability, for now we use the two terms interchangeably.) We first calculate the weighted fraction of people aged 50-64 with self-reported health that is either fair or poor, across all European countries in the SHARE sample and the U.S.[9] We graph these percentages against the fraction of people age 50-64 who are receiving DI benefits (as in Table 1), with results shown in Figure 4. While Sweden, in the lower-right hand corner of the graph, appears to be consistent with our basic hypothesis – with nearly fifteen percent enrolled in DI and fewer than 40 percent reporting being in fair and poor health -- Denmark registers an even higher percentage on DI (15.7), but with 65 percent of the DI population claiming fair or poor health. The Netherlands have a similar fraction of the DI population claiming fair or poor health, but they stand out with more than 17 percent of 50-64 individuals receiving DI benefits. The overall correlation is faintly negative (largely because of Sweden), but is not significant. Nor is there any association between depression and the percentage enrollment for the DI program.

Table 1 also shows the percentage of people age 50-64 reporting fair or poor health by country for the DI and non-DI population. There is considerable variation across countries in this percentage, for both those not receiving DI and those receiving it. For example, 29 percent of non-DI recipients in Spain and 26 percent in Germany report fair or poor health, in contrast to just 5 percent in Sweden and 9 percent in Switzerland. For those receiving DI, there is also remarkable variation, again from 38 percent in Sweden to 76 percent in the US and 80 percent in Germany. We have argued that the difference in these measures can at least (in theory) remove the additive mean country effects, and so the third column in Table 1 shows these differences.

[9] We focus on the 50-64 population because in the US as well as in several European countries many people make the transition from disability insurance to the old-age Social Security program at age 65.

There is far less variation overall, and somewhat different patterns, but Germany, Belgium, Greece, the Netherlands, Denmark, and the U.S. all demonstrate differences of 50 percent or more. The correlation between the difference and the percentage of the population on DI is essentially zero. Finally, Table 1 shows a somewhat different statistic: the percentage of people who report either fair or poor health who are covered under the DI system.[10] This ranges considerably depending on the coverage of the program, from 8 percent in Austria, and 11 percent in Switzerland and Greece, to 57 percent in Sweden. To the extent that self-reported health is a good proxy for ability to work, the mirror-image of this measure (100 minus this percentage) is a rough measure of the "Type I" error of people who are disabled but are not covered by DI.

Somewhat different results are found for the CES-D depression score. Again, there are large variations across countries in the absolute level of depression but less variation in the differences. Denmark as well exhibits less of a gradient in depression by DI status compared to its previous pattern for self-reported health. The US and Switzerland stand out as the countries with the largest differential in terms of people with self-reported symptoms associated with depression.

Finally, we consider probit estimates of the model, shown in Table 3. Consider first the regressions for the European countries (pooled) and the US. (The results are similar when dummy variables are included for each country.) The largest difference between the two specifications is that education (or market opportunities) is associated with the likelihood of being on disability in European countries conditional on health status, but this does not hold in the U.S. (When not controlling for self-reported health, however, education is strongly

[10] Equal to the number of people in DI who say they are in fair or poor health divided by the universe of people who are in fair or poor health.

associated with the probability of being on DI in the U.S.) Considering the ratio of the two marginal effects suggests that the U.S. line looks more like nn' in Figure 4, compared to mm' for European countries. Separate regressions are provided for each country in the remainder of Table 3; there is variation in these estimates, although most of the education coefficients are larger than those found in the U.S. Thus the U.S. appears to follow the "medicalization" model relying (however imperfectly) on health rather than labor market opportunities.

Consider next the coefficients for Denmark (DK) and Sweden (SE). Both countries experience similar enrollment rates. We can interpret the coefficient on tertiary education as a measure of how important are market opportunities, and the coefficient on poor health as a marker for how important is health. Recall from the theory section that the ratio of the coefficient for poor health (.145 for Sweden, .452 for Denmark) divided by the coefficient for tertiary education (-.029 in Sweden, -.096 in Denmark) is a measure of the relative weight placed on health compared to market opportunities. These ratios are very similar in the two countries (20 and 21), suggesting that both countries view their program as insurance against the risk of poor job prospects. The magnitude of the Danish coefficients, however, are three times the rates for Sweden, suggesting that the Danish system does a better job of screening applicants. While anecdotal, a recent OECD study (2009) states:

> *Sweden* is a nation with an historically strong ethos of social protection and it is seeking to tackle the capacity assessment challenge through a Work Capacity Commission tasked with receiving submissions and providing a forum for public discussion.
>
> The disability scheme in *Denmark* which was reformed in 2003 incorporates a most fundamental conceptual shift. Disability assessment is now focused on what a person can do rather than their loss of capacity; more precisely, the extent to which a person is able to carry out a subsidised job (a so-called "flex-job"). A disability benefit is only granted where capacity is held to be permanently reduced to the extent that a flex-job cannot be performed, and participation in rehabilitation would

> not help to restore this capacity. In determining capacity, a comprehensive individual resource profile is being put together which includes measures of health, social and labour market proximity criteria. In this respect, Denmark is a best-practice example within the OECD (p. 19).

At this stage, we have not developed a summary statistic to explain how well the model fits the overall data. While the Netherlands, for example, shows very high rates of utilization, the relative and absolute weight placed on health appears to be very high. By contrast, France, exhibits roughly average European coefficients, which does not easily explain why its rate of enrollment is so low.

5. Conclusion

In this paper, we study the reported health and work opportunities of the population age 50-64 enrolled in a disability insurance (DI) program across a sample of 11 European countries and the United States. While these results are still preliminary, we can suggest three basic results arising from our theoretical and empirical analysis.

First, there appears to be little or no association between the enrollment rates for DI programs and the average self-reported health of the DI population. This seems somewhat puzzling, in that one might expect a tension in most programs between extending coverage and enrolling successively less disabled people as a result. This result holds even when we adopt a more robust measure (relative to differences across countries in underlying health status), which is the difference between measured health status for the DI and the non-DI populations.

Second, we develop a model in which these aggregate patterns make sense given that different countries could place different weights on the two general criteria for eligibility – health status and wage opportunities or productivity at work. We use education as a largely exogenous proxy for market opportunities, and show that differences in the weights that individual countries

(and individuals in those countries) place on health (α) and market opportunities (γ) can in theory explain the empirical puzzles we observe.

Third, we use the individual SHARE and HRS data to show that differences in α and γ across countries may help to explain the aggregate puzzle. The U.S., for example, exhibits modestly high rates of enrollment, but its DI enrollees appear to exhibit the poorest health (whether overall health or depression) compared to any other country in the sample. One reason is that, *conditional* on self-reported health, educational attainment has no impact in the U.S. of the likelihood of DI enrollment; thus in the US one does not find relatively more healthy people with worse market opportunities. By contrast, in most European countries lower educational attainment leads to a much higher chance of being on DI, even controlling for health status.

The U.S. also exhibits the highest rates of rejections for disability applicants (OECD, 2009). This may in part be the consequence of a strict "medical model" of DI eligibility – that it's not enough to be less productive at work, one also needs an established and severe clinical disease to gain eligibility. And the US DI program makes it very difficult indeed to qualify for DI, with recipients who report broken bones and chronic diseases often waiting 4 or 5 years for final decisions, during which they're not supposed to be working (Eckholm, 2007). Thus the dependence of enrollment on health status is certainly consistent with (a) a system that depends nearly entirely on poor health for eligibility, but which in turn (b) draws from a population in fairly poor health to begin with.

Furthermore, there is at least anecdotal evidence that musculoskeletal pain and depression – both very important sources of disability – are affected by incentives for being rewarded for the presence of such pain. One example comes from Cassidy et al (2000) who reported a quicker resolution to whiplash injury following a change from tort liability to no-fault

liability in Canada. Of course, there are strong incentives of patients still in litigation to report injuries strategically. However, Cassidy (private communication) suggested better outcomes for the no-fault population after the cases were resolved and there was no longer an incentive to bargain over settlements.

Similarly, an epidemic of repetition strain injury (RSI) in Australia suggests that social norms helped the debilitating ailment spread like an infectious disease; some factories or states were affected in large numbers and others not (Gawande, 2002). The epidemic ended once the syndrome was no longer in favor with physicians, and after disability eligibility for RSI was tightened by the Australian government. Rege et al. (2008) present evidence in favor of this type of social contagion model, while Hadler, Tait, and Chibnall (2007) proposed that chronic back pain often evolves in response to the strong incentives inherent in the U.S. workers' compensation system.

The pessimistic view of these results would emphasize the large differences in the structures of DI programs across countries that could reflect inefficiencies; these are differences that are the result of policy factors and not the health of the population (Borsch-Supan, 2007; Börsch-Supan and Roth, 2010). The optimistic view of these country-level variations is that some countries may have developed better methods of avoiding the inherent tradeoffs between Type 1 and Type 2 error. (Certainly over time countries are evolving, for example in the recent reforms in the Netherlands that have reduced rates of DI enrollment substantially since 2004.)

The most promising approach appears to be to divert workers away from the medical model of disability and to remain working despite the disability or pain. The recent successes of supportive employment in the U.S. and Europe, by which mentally disabled people are encouraged to return to appropriate work, shows considerable promise (Burns et al, 2007), and

there is at least suggestive evidence that at least in the U.S., such programs can pay for themselves by reducing disability and medical costs (Drake, et al., 2009).

References

Apkarian, A. Vania, Marwan N. Baliki, and Paul Y. Geha, "Toward a Theory of Chronic Pain," *Progress in Neurobiology* 87 2009: 81-97.

Atlas, Steven J., RA Deyo, KB Keller, et al. "The Maine Lumbar Spine Study, Part II. 1-year outcomes of surgical and nonsurgical management of sciatica," *Spine* 21, 1996: 1777-86.

Atlas, Steven J, Chang Y, Keller RB, Singer DE, Wu YA, Deyo RA, "The impact of disability compensation on long-term treatment outcomes of patients with sciatica due to a lumbar disc herniation," *Spine* 31, 2006: 3061-9.

Atlas, Steven J., and Jonathan Skinner, "Education and the Prevalence of Pain," NBER Working Paper (submitted, April 2009).

Autor, David H. and Mark G. Duggan, "The Rise in the Disability Rolls and the Decline in Unemployment." *Quarterly Journal of Economics,* 118:1, February 2003: 157–206.

Autor, David H. and Mark G. Duggan, "The Growth in the Disability Insurance Roles: A Fiscal Crisis Unfolding," *Journal of Economic Perspectives* 20(3), Fall 2006: 71-96

Banks, James, Arie Kapteyn, James P. Smith, and Arthur van Soest, " International Comparisons of Work Disability", RAND Working Paper, WR-155, 2004.

Banks, James, Michael Marmot, Zoe Oldfield, and James P. Smith, "Disease and Disadvantage in the United States and in England," *JAMA* 295(7) May 3, 2006: 2037-2045.

Benitez-Silva, Hugo, Buchinsky, Moshe, Chan, Hiu Man, Cheidvasser, Sofia and Rust, John P., "How Large is the Bias is Self-Reported Disability?" *Journal of Applied Econometrics* 19, 2004: 649-670.

Börsch-Supan, Axel, "Work Disability: The Effects of Demography, Health, and Disability Insurance," Working Paper, University of Mannheim, October 2007.

Börsch-Supan, Axel, "Changes in Health Status and Work Disability," in A. Börsch-Supan, et al. (eds.), Health, Ageing, and Retirement in Europe – Starting the Longitudinal Dimension. Mannheim, pp.228-236.

Börsch-Supan, Axel, and Henning Roth, "Work Disability and Health over the Life Course," Working Paper, University of Mannheim, May 2010.

Burns, Tom, Jocelyn Catty, Thomas Becker, Robert E. Drake, Angelo Fioritti, et al., "The Effectiveness of Supported Employment for People with Severe Mental Illness: A Randomised Controlled Trial," *The Lancet*, 370(9593), 29 Sept/5Oct 2007:1146-52.

Case, Anne, and Angus Deaton, "Broken Down by Work and Sex: How Our Health Declines," in D. Wise (ed.) Analyses in the Economics of Aging, University of Chicago Press and NBER, 2005.

Cassidy, J. David, Linda J. Carroll, Pierre Côté, Mark Lemstra, Anita Berglund, and Åke Nygren,, "Effect of Eliminating Compensation for Pain and Suffering on the Outcome of Insurance Claims for Whiplash Injury," *NEJM* 342, April 20, 2000: 1179-1186.

Cutler, David, and Adriana Lleras-Muney, "Education and Health: Evaluating Theory and Evidence," NBER Working Paper No. 12352, July 2006.

Deyo, R. and Y. Tsui-Wu, "Functional Disability due to Back Pain: A Population-based Study Indicating the Importance of Socioeconomic Factors," *Arthritis and Rheumatism* 2005. **30**: 1247-1253.

Drake, Robert E., Jonathan S. Skinner, Gary R. Bond, and Howard Goldman, "Social Security Disability: Making it a Safety Net for Supportive Employment," *Health Affairs*, 2009 (forthcoming).

Eckholm, Erik, "Disability Cases Last Longer as Backlog Rises," *The New York Times*, December 10, 2007.

European Commission, *Adequate and Sustainable Pensions*, Joint Report by the Commission and the Council, Luxembourg: European Communities, 2006.

European Commission, *MISSOC Tables 2007*, Luxembourg: European Communities, 2007.

Social Protection Committee, *Promoting Longer working Lives through Pension Reforms. Second Part. Early Exits from the Labour Market*, European Commission, Directorate-General for Employment, Social Affairs and Equal Opportunities, 2008.

Euwals, Rob, Ruud de Mooij, and Daniel Van Vuuren, "Rethinking Retirement: From Participation towards Allocation," CPB Netherlands Bureau for Economic Policy Analyis No. 80, April 2009.

Gawande, Atul, *Complications.* New York: Henry Holt and Company, 2002.

Glaeser, Edward, Bruce Sacerdote, and Jose Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics*, 111, 1996: 507-548.

Glaeser, Edward., Bruce Sacerdote, and Jose Scheinkman, "The Social Multiplier," *Journal of the European Economic Association*, April-May 2003.

Hagen, K., et al., "Socioeconomic Factors and Disability Retirement from Back Pain: A Population-Based Prospective Study in Norway 1983-1993," *Spine*, 2000.

Jensen, Maureen C., Michael N. Brant-Zawadzki, Nancy Obuchowski, Michael T. Modic, Dennis Malkasian, and Jeffrey S. Ross, "Magnetic Resonance Imaging of the Lumbar Spine in People without Back Pain," *New England Journal of Medicine*, 331(2), July 14, 1994: 69-73.

Kapteyn, Arie, James Smith, and Arthur van Soest, "Dynamics of Work Disability and Pain," Rand Working Paper WR-387, March 2006.

Kapteyn, A., J. P. Smith, and A. van Soest," Work Disability, Work, and Justification Bias in Europe and the U.S.," NBER Working Paper 15245, August 2009.

Krause, N., et al., "Psychosocial Job Factors, Physical Workload, and Incidence of Work-Related Spinal Injury: A 5-Year Prospective Study of Urban Transit Operators," *Spine*, **23**(23), 1998: 2507-2516.

Krueger, Alan B., and Arthur A. Stone, "Assessment of pain: a community-based diary survey in the USA," *The Lancet*, 371(9623), May 3, 2008: 1519-1525.

McVicar, Duncan, "Why Have UK Disability Benefit Rolls Grown So Much," Journal of Economic Surveys, 22(1), 2008: 114-39.

Meijer, Erik, Arie Kapteyn, and Tatiana Andreyeva," Health Indexes and Retirement Modeling in International Comparisons," RAND Working Paper WR-614, August 2008.

Melzack, Ronald, "Pain: Past, Present, and Future," *Canadian Journal of Experimental Psychology* 47(4) 1993: 615-629.

Meng, Chun-Lo, and Peter Schmidt, "On the Cost of Partial Observability in the Bivariate Probit Model," *International Economic Review* 26(1), February 1985: 71-85.

MISSOC Secretariat, *MISSOC Analysis 2008. Social Protection: Aspects Of Flexicurity And Active Inclusion*, European Commission, DG Employment, Social Affairs & Equal Opportunities, June 2008.

OECD, *Transforming Disability into Ability*, Paris: OECD, 2003.

OECD, "Sickness, Disability, and Work: Keeping on Track in the Economic Downturn," High-Level Forum, Stockholm, 14-15 May 2009.

Poirier, Dale J., "Partial Observability in Bivariate Probit Models," *Journal of Econometrics* 12, 1980: 209-217.

Rege, Mari, Kjetil Telle, and Mark Vortruba, "Social Interaction Effects in Disability Pension Participation: Evidence from Plant Downsizing," working paper, Case Western Reserve University, June 2008.

Figure 1: Graph Showing Distribution of Individuals by Health and Wage Rate (or Market Opportunities), and the influence of Disability Insurance

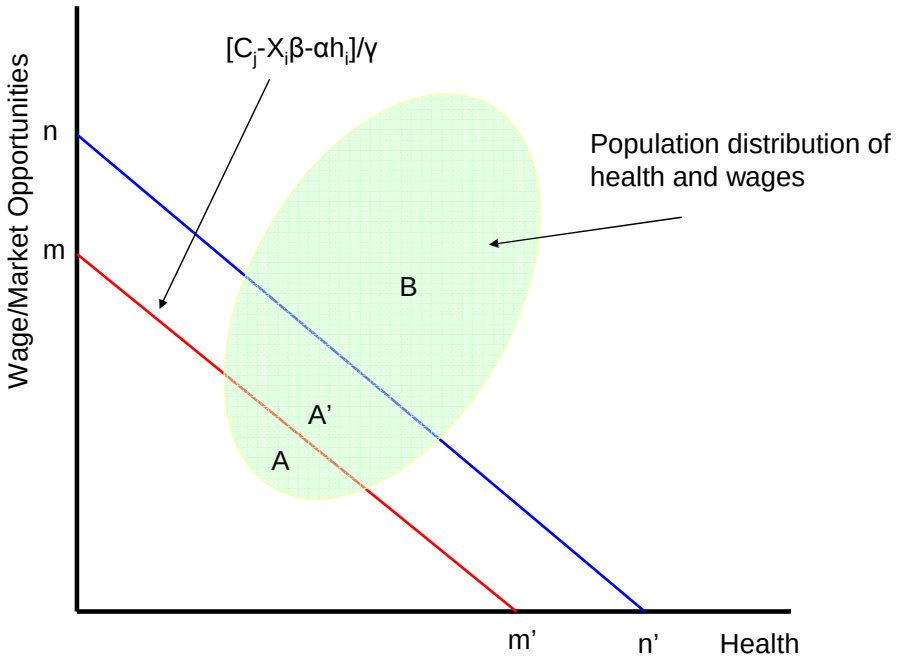


Figure 2: Graph Showing Distribution of Individuals by Health and Wage Rate (or Market Opportunities) for Two Countries, R and P

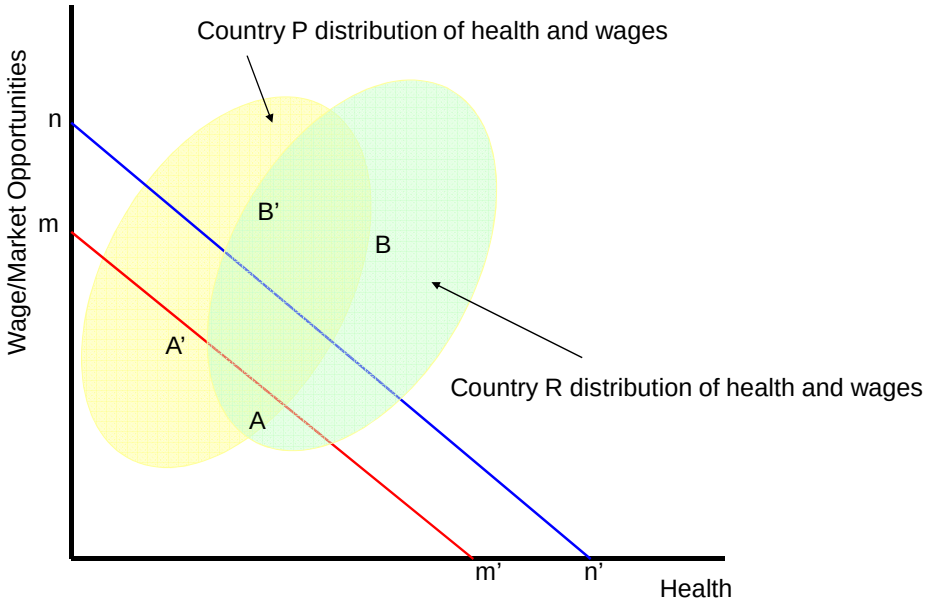


Figure 3: Graph Showing Distribution of Individuals by Health and Wage Rate (or Market Opportunities), and Different Country-Level Disability Systems

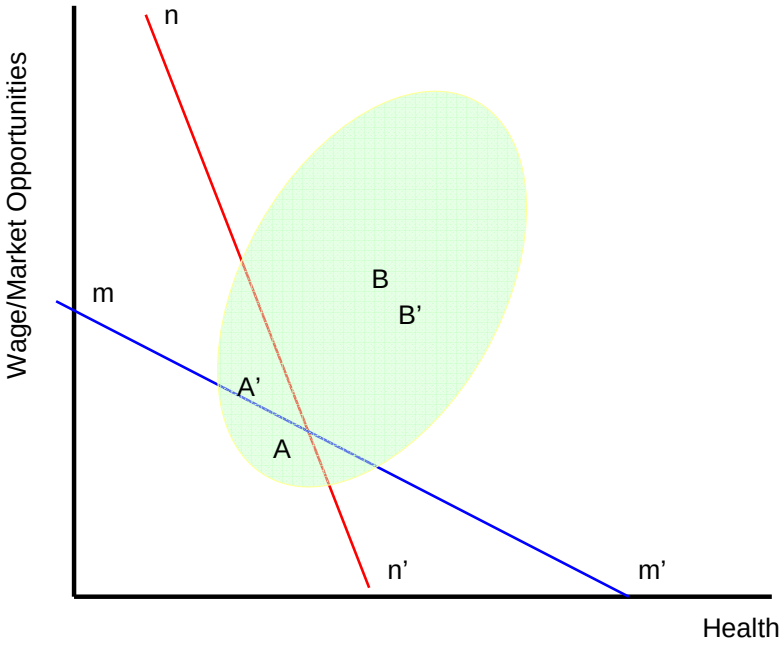
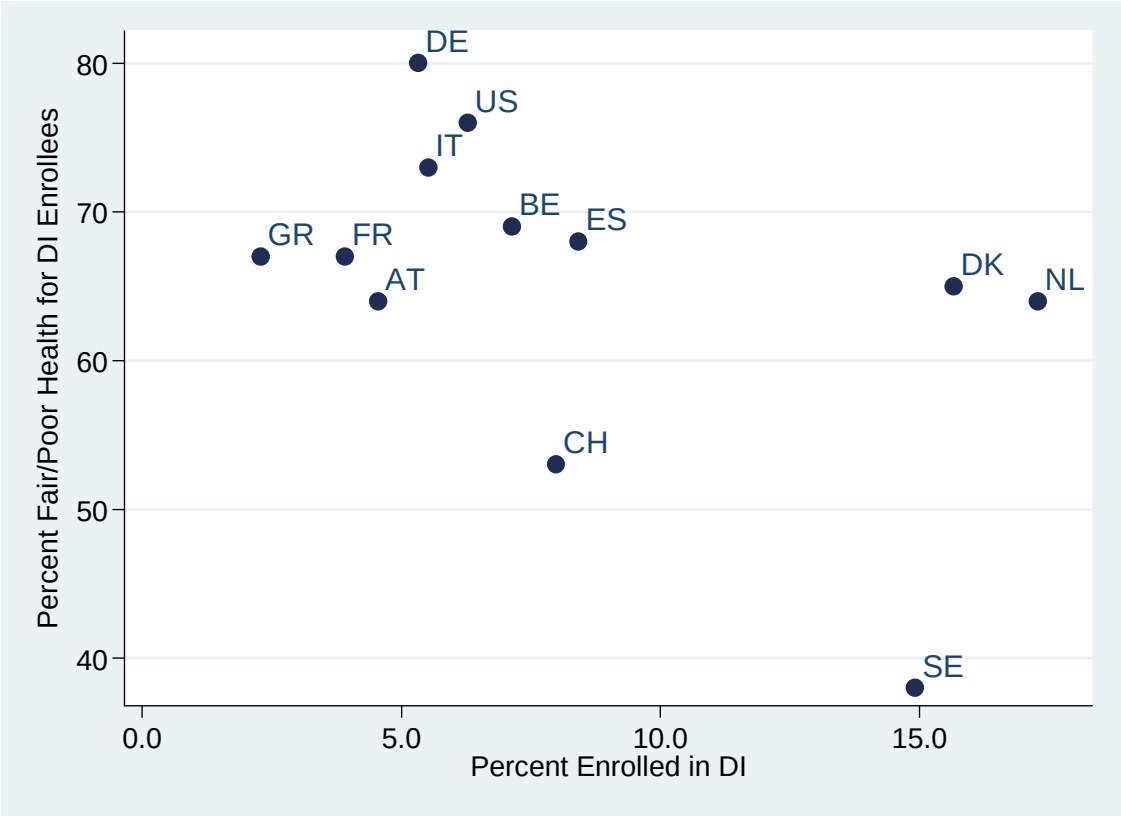
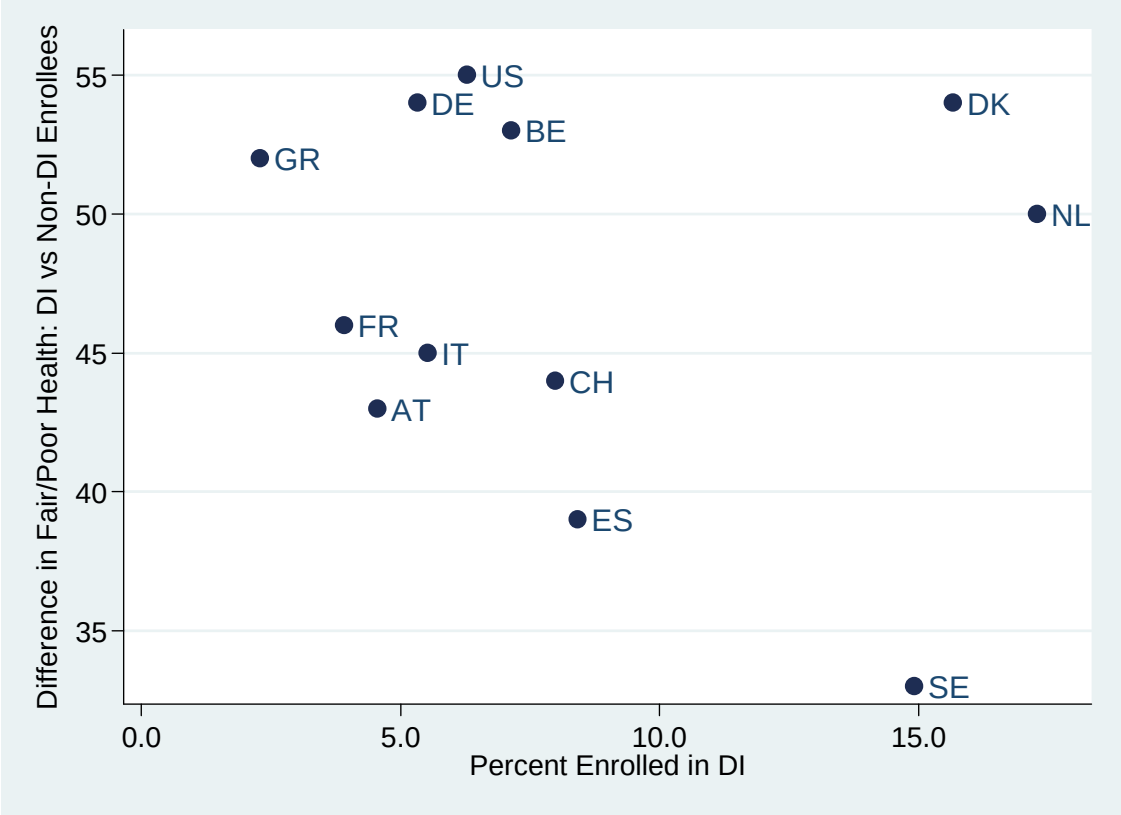


Figure 4: Scatter Diagram of the Percentage of Self-Reported Fair/Poor Health for Disability Insurance Enrollees, 12 Countries



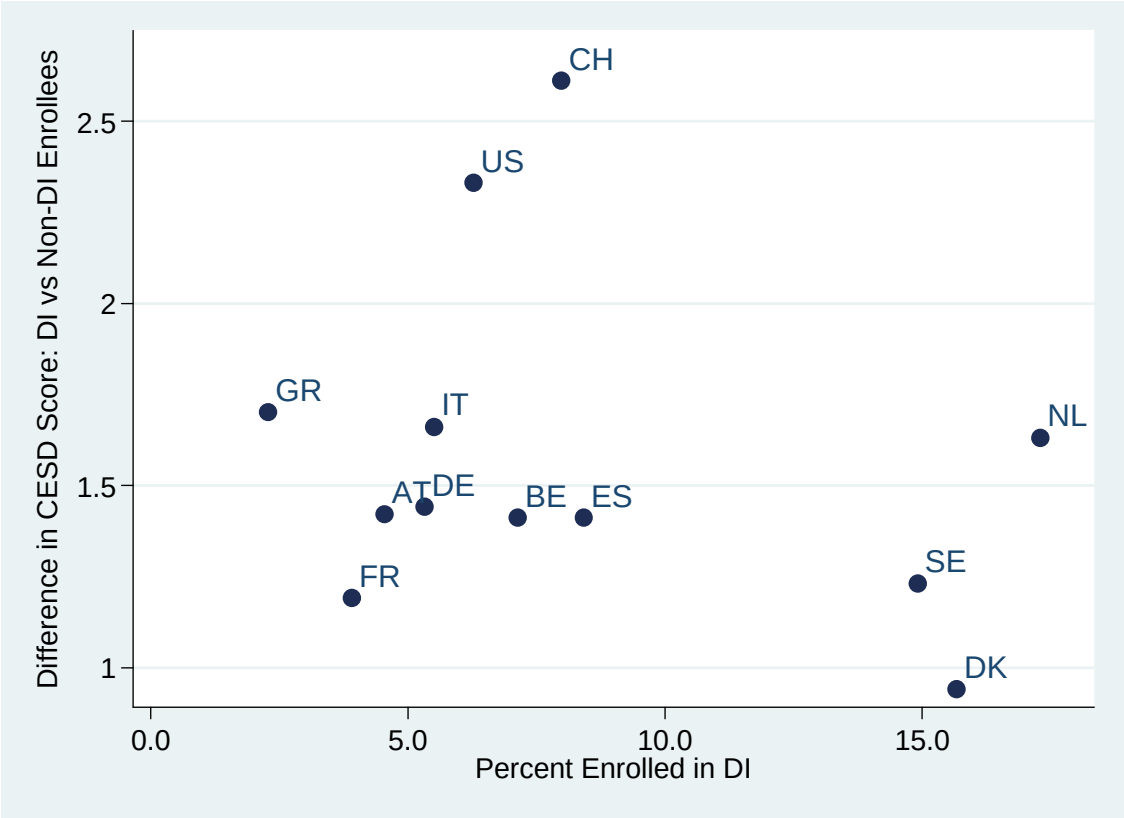
Source: Authors' calculations using SHARE 2004 and HRS 2004. Age 50-64. Population-weighted data.

Figure 5: Scatter Diagram of the Difference (Between DI and non-DI Enrollees) in the Percentage Reporting Fair/Poor Health for Disability Insurance Enrollees, 12 Countries



Source: Authors' calculations using SHARE 2004 and HRS 2004. Age 50-64. Population-weighted data.

Figure 6: Scatter Diagram of the Difference (Between DI and non-DI Enrollees) in CESD Depression Score, 12 Countries



Source: Authors' calculations using SHARE 2004 and HRS 2004. Age 50-64. Population-weighted data.

Table 1: Percentage of People in Fair or Poor Self-Reported Health, by Disability Insurance Status and Country

	Not on Disability Insurance	Receiving Disability Insurance	Difference	% on DI	% Fair/poor Enrolled in DI
United States	21	76	55	6.3	20
Sweden	5	38	33	14.9	57
Denmark	11	65	54	15.7	53
Germany	26	80	54	5.3	18
Netherlands	14	64	50	17.3	40
Belgium	16	69	53	7.1	19
France	21	67	46	3.9	11
Switzerland	9	53	44	8.0	30
Austria	21	64	43	4.6	08
Italy	28	73	45	5.5	13
Spain	29	68	39	8.4	16
Greece	15	67	52	2.3	11

Percentage of individuals not receiving disability insurance and self reporting being in poor or fair health (column 1), receiving disability insurance and self reporting being in poor or fair health (column 2), receiving disability insurance (column 4), self reporting being in poor or fair health and receiving disability insurance (column 5). Column 3 contains the difference between the first two columns.
Source: Authors' calculations using SHARE 2004 and HRS 2004. Age 50-64. Population-weighted data.

Table 2: CESD Depression Scores, By Disability Insurance Status and Country

	Not on Disability Insurance	Receiving Disability Insurance	Difference	% on DI
United States*	-1.18	1.15	2.33	6.3
Sweden	-1.78	-0.55	1.23	14.9
Denmark	-2.23	-1.29	0.94	15.7
Germany	-1.62	-0.18	1.44	5.3
Netherlands	-2.11	-0.48	1.63	17.3
Belgium	-1.61	-0.20	1.41	7.1
France	-1.32	-0.13	1.19	3.9
Switzerland	-2.01	0.60	2.61	8.0
Austria	-1.64	-0.22	1.42	4.6
Italy	-0.66	1.00	1.66	5.5
Spain	-1.03	0.38	1.41	8.4
Greece	-1.05	0.65	1.70	2.3

Percentage of individuals not receiving disability insurance and self reporting being depressed (column 1), receiving disability insurance and self reporting being depressed (column 2), receiving disability insurance (column 3), self reporting being depressed and receiving disability insurance (column 4).
* One point for every "yes" or "no" denoting less depression. All other countries: Almost all of the time or most of the time denotes yes, some of the time, and almost none of the time denotes no.
Source: Authors' calculations using SHARE 2004 and HRS 2004. Age 50-64. Population-weighted data.

Table 3: Probit Regression Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Euro.	US	SE	DK	DE	NL	BE	FR	CH	AT	IT	ES	GR
Respondent is male	0.011	0.013	-0.056	-0.031	0.020	0.043	0.034	0.011	0.010	0.025	0.023	0.048	0.002
	(0.004)	(0.004)	(0.013)	(0.020)	(0.009)	(0.016)	(0.008)	(0.007)	(0.011)	(0.012)	(0.011)	(0.016)	(0.006)
Age 55 – 59	0.005	0.001	0.017	-0.002	0.027	0.022	-0.006	-0.004	0.006	-0.01	0.002	-0.011	-0.001
	(0.005)	(0.005)	(0.018)	(0.024)	(0.014)	(0.019)	(0.007)	(0.006)	(0.015)	(0.013)	(0.013)	(0.015)	(0.006)
Age 60 – 64	-0.001	-0.01	0.022	-0.063	0.011	0.113	-0.005	-0.026	0.019	-0.037	-0.012	-0.007	0.001
	(0.005)	(0.004)	(0.020)	(0.026)	(0.013)	(0.026)	(0.009)	(0.008)	(0.018)	(0.016)	(0.014)	(0.015)	(0.007)
Married	-0.046	-0.024	-0.06	-0.126	-0.018	-0.149	-0.036	-0.031	-0.069	-0.029	-0.013	-0.075	0.007
	(0.006)	(0.005)	(0.023)	(0.033)	(0.012)	(0.033)	(0.013)	(0.011)	(0.030)	(0.016)	(0.015)	(0.025)	(0.005)
Secondary education	-0.022	-0.004	-0.023	-0.085	0.006	-0.013	-0.016	-0.007	-0.027	-0.005	-0.024	-0.043	-0.010
	(0.004)	(0.005)	(0.014)	(0.024)	(0.011)	(0.018)	(0.006)	(0.006)	(0.009)	(0.012)	(0.009)	(0.011)	(0.005)
Tertiary + education	-0.023	-0.004	-0.029	-0.096	-0.012	-0.013	-0.017	-0.009	-0.023	-0.007	-0.028	-0.047	-0.009
	(0.004)	(0.005)	(0.015)	(0.022)	(0.012)	(0.018)	(0.007)	(0.007)	(0.012)	(0.015)	(0.011)	(0.011)	(0.005)
Retired	0.026	0.130	0.378	0.250	0.002	-0.108	-0.035	-0.004	-0.023	0.036	0.018	0.048	0.058
	(0.006)	(0.014)	(0.038)	(0.053)	(0.012)	(0.011)	(0.007)	(0.010)	(0.011)	(0.015)	(0.012)	(0.027)	(0.021)
Excellent health	-0.047	-0.034	-0.098	-0.098		-0.090	-0.031	-0.018	-0.036		0	-0.009	
	(0.004)	(0.004)	(0.012)	(0.019)	b	(0.015)	(0.006)	(0.007)	(0.013)		(0.021)	(0.029)	
Very good health	-0.044	-0.035	-0.089	-0.067	-0.027	-0.075	-0.018	-0.014	-0.052	-0.024	0	-0.014	
	(0.004)	(0.004)	(0.013)	(0.022)	(0.010)	(0.016)	(0.008)	(0.007)	(0.014)	(0.012)	(0.018)	(0.017)	
Fair health status	0.098	0.052	0.118	0.195	0.078	0.273	0.097	0.048	0.09	0.03	0.078	0.083	0.048
	(0.008)	(0.010)	(0.039)	(0.052)	(0.019)	(0.032)	(0.022)	(0.017)	(0.049)	(0.020)	(0.022)	(0.026)	(0.018)
Poor health status	0.277	0.133	0.145	0.452	0.203	0.535	0.379	0.185	0.197	0.151	0.318	0.271	0.212
	(0.020)	(0.020)	(0.062)	(0.082)	(0.053)	(0.074)	(0.067)	(0.051)	(0.106)	(0.063)	(0.066)	(0.062)	(0.086)
Mean of dependent variable	0.082	(0.078)	0.150	0.164	0.055	0.158	0.067	0.040	0.080	0.048	0.056	0.082	0.037
N of obs	14530	7942	1588	909	1446	1689	1938	1565	500bb	836	1331	1071	727

Source: Authors' calculations using SHARE 2004 and HRS 2004. Age 50-64. Robust Standard Errors in parentheses.