

**Mandatory Disability Insurance Systems: What are the Issues, What are Countries
Doing and How can they do Better?***

By Estelle James

Prepared for the World Bank, February 2011

(Draft paper, not for quotation)

*I wish to thanks Christopher Prinz, head of the OECD disability project, for his generous help, and Andrew Reilly, OECD, for assistance with access to the OECD Social Expenditure Database (SOCX).

Contents

- I. An Overview of Traditional Disability Insurance Schemes
 - Defined benefits
 - Spending and reciprocity rates
 - Distribution of disability benefits
 - Diagnoses
 - Relationship to unemployment benefits and other social insurance programs
- II. Problems—1) Disability is non-observable and subjective, hence classification errors
 - Classification errors based on audits
 - Different approval rates among examiners, administrative judges and localities
 - System decisions vs. self-reported health status
 - System decisions vs. mortality rates and body mass
 - The trade-off between errors of inclusion and exclusion
- III. Problems—2) Moral hazard, agency problems and social norms
 - Moral hazard
 - Agency problems
 - The role of social norms and contagion
 - Implications for tagging the disabled
- IV. Problems—3) High costs and reciprocity rates: sensitivity to benefit size and screening
 - Costs of disability insurance
 - Why the divergence across countries and changes over time?
 - Disability applications respond to benefit levels and replacement rates
 - Applications and enrollments respond to screening procedures
 - Disability reciprocity rate varies with unemployment rate
 - Demand for disability benefits depends on access to other social insurance (and v.v.)
 - Relationship to moral hazard and non-observability of disability
- V. Problems—4) Low labor force participation rates of the disabled
 - How does the disability system affect labor force participation (LFP)?
 - Disability benefit amount and replacement rate affects LFP: U.S. studies
 - Disability benefit amount affects LFP: Canada
 - Tightness of screening affects work probabilities
 - The impact of recent reforms
- VI. Reforms—1) Changing the system's parameters
 - Eligibility
 - Is eligibility for contributors only or also for non-contributors?
 - Is eligibility based on incapacity to do one's regular job or any job?
 - Are the partially disabled eligible for benefits?
 - Benefit amount and replacement rate
 - Screening
 - How stringent is the screening mechanism?
 - Sources of medical information—individuals' own GP or system's physicians?
 - Who else participates in the screening procedure?
 - How to encourage screening by self-selection

Sick leave and anti-selection

Is certification permanent? Is periodic reassessment required?

Should the disabled be treated differently from other unemployed individuals?

Some examples

Poland

Sweden

VII. Reforms—2) Increasing work incentives

Thresholds

Trial work periods

Eliminating the implicit tax

Positive rewards for work—wage subsidies and supplements

Anti-discrimination legislation, work place accommodation and employment quotas

VIII. Structural Reforms that Improve Information and Monitoring: A Larger Private Sector Role

The Netherlands: employer responsibility for financing and monitoring

Contrast with Denmark—subsidized flex-jobs by municipalities

Australia—outcome-based payments to service providers

Chile—countervailing information from pension funds and insurance companies

IX Conclusion

Tables

Table 1: Characteristics of Traditional Disability Schemes and Recent Reforms

Table 2: Key Policy Design Issues

Table 3: Public and Mandatory Private Spending on Incapacity as % of GDP (compared with old age spending), OECD Countries, 2005

Table 4: Outcomes to be Monitored

Figures

Figure 1: Spending on incapacity as a % of GDP, 1995, 2000, 2008

Figure 2: Public + Mandatory Private Spending on Disability/Old Age Spending, 2005

Figure 3: Disability benefit recipients as % of population age 20-64, 1995 and 2008

Figure 4: Outflow Rates as % of All Disability Beneficiaries, mid 2000's

Figure 5: Inflow to disability benefits per 1000 working age population, 2000 and 2008

References

Endnotes

Mandatory Disability Insurance Systems: What are the Issues, What are Countries Doing and How can they do Better?

Mandatory old age insurance exists to make sure that people have some source of income when they become too old to work productively. It protects against a severe drop in standard of living among the elderly due to myopia and insufficient saving when young or unexpectedly high longevity when old. Saving is a feasible part of an old age security program because people work for many years before retirement and they face a high probability of reaching old age.

However, a substantial number of people lose productive capacity before they grow old because they have a major health problem that makes it difficult for them to work at their normal job, or at any job. According to household surveys in OECD countries, 14% of the working age population regard themselves as having a health problem that limits their daily activities, such as work (OECD 2010). These individuals and their families face a potentially large loss of earning power and standard of living. This is the *raison d'être* for disability insurance. It guards against a small risk of a very large loss—the situation for which insurance is particularly suited. In addition, the disabled face higher living costs due to the need for supportive goods and services, which disability insurance can help to cover.¹

Why not leave this insurance to the voluntary market? First, people may be myopic -- have a short time horizon, underestimate the risk of becoming disabled, and hence don't purchase enough insurance voluntarily. Second, those who expect to become disabled, perhaps because they already notice some signs of bad health, are most likely to purchase disability insurance. Then, the risk associated with the average insurance purchase is higher than the average risk in the over-all population. Insurance companies build this "adverse selection" of consumers into their pricing policies—making the price "too high" for the average potential consumer and further discouraging him from purchasing disability insurance voluntarily. A third obstacle in voluntary markets stems from the attempts by insurance companies to "cherry-pick" the better (lower) risks. As discussed below, disability is largely an unobservable state at the individual level. However, it is easier to predict who has the characteristics of a high-risk group—for example, older workers and those in hazardous occupations. Members of these groups may face high prices that keep them out of the market, even if the individual is healthy.

Many countries make disability insurance mandatory to avoid these problems of myopia, adverse selection and cherry picking based on group characteristics. Insurance benefits maintain the living standard of the disabled and their families and prevent poverty due to severe health problems that impede earning capacity before old age. Chandra and Samwick (2008) estimate that an average rational non-myopic worker would be willing to pay 5% of expected consumption to eliminate the uncertainty associated with disability risk.² This is not far from the average percentage of payroll spent on disability, in OECD countries.

However, insurance creates other problems, and in mandatory systems these problems are magnified because of their broad coverage. The classic problems are: *moral hazard*—insurance increases risky behavior that leads a person to qualify for benefits; *high costs* that are due, in part, to moral hazard; and *reduced labor supply*, hence GDP, as individuals who could have worked instead qualify for disability benefits so need not work. Indeed, they are often discouraged from work by system incentives.

Underlying all these problems is *the difficulty in objectively assessing* whether or not a person is truly disabled and whether a medical problem makes work impossible, hence the *frequency of errors*. In countries where this has been studied, one-third or more of all decisions about claims are estimated to be wrong—either false negatives or false positives. If an objective test existed, moral hazard would be reduced since people would not qualify for benefits unless they were truly disabled and the rest would have to work to consume. But in the absence of an objective test, these problems of moral hazard, high costs and reduced labor supply abound. The difficulty in assessing functional disability has strong implications for the optimal degree of screening and benefit generosity, as well as the desirability of treating the disabled differently from other social insurance recipients such as the unemployed or early retirees. It also limits the kind of disability program that is feasible in low income countries skilled human capital is scarce and work incapacity due to lack of education or unemployment are plentiful.

Part I of this paper presents an overview of traditional disability benefit schemes—the prevalence of disability, recipiency and spending rates, the distribution of benefits, the kind of diagnoses that have led individuals to receive them and the relationship to other social insurance programs. Parts II, III, IV and V elaborate on problems that have developed—classification errors, moral hazard, high costs and reduced labor supply—from the viewpoint of economic theory and empirical tests. Analyses show that exogenous variables such as unemployment rate

and population aging play a role in explaining differences in costs and labor force participation across countries and time, but policy variables are most important.

Parts VI, VII and VIII review recent reforms that seek to reduce moral hazard and costs while raising accuracy and work. Some of these reforms are parametric, changing procedures within the same basic structure--e.g. tightening eligibility constraints, improving information and screening methods, requiring periodic reassessment and bringing greater symmetry between disability and other forms of social insurance (Part VI). While reducing benefit amounts would seem to be the most direct parametric change, countries have been reluctant to use it because it also cuts the insurance value. Instead, the most interesting reforms operate through improved *information and incentives*.

For example, in some reformed systems the group that is totally unable to work is separated from the group that retains some work capacity and the latter is given subsidies for work instead of benefits for not working (Part VII). In more dramatic reforms, informational and financing responsibilities are shifted away from public agencies toward private sector employers, pension funds, insurance companies and service providers, who have or are given a pecuniary incentive to cut costs while moving the disabled back to work (Part VIII). The object is to retain welfare-enhancing insurance benefits while minimizing their distortionary effects. The jury is still out on how effective these structural reforms will be. Tables 1 and 2 summarize key design questions, Table 3 gives spending on incapacity in many OECD countries, Table 4 lists outcomes that should be regularly monitored and Figures 1-5 display some of these outcomes.

Many people who self-report themselves as having medical impairments that hamper their work or other daily activities do not receive disability benefits. And some who already receive benefits do not self-report themselves as disabled. In this paper we focus on “recipients,” although we do occasionally refer to self-reported disability as a point of comparison. We are concerned about features of the system that enable both groups to achieve a reasonable standard of living without a high taxpayer cost, by utilizing their remaining capacities to work productively. Recent reforms have the object of keeping the self-reported disabled off the future disability benefit rolls and in the labor force. This may also prevent their disability condition from getting worse. Most of the evidence in this paper is from high-and middle-income countries, where data are available. Wherever possible, we infer how the situation and recommended policies may differ in low and middle-income countries.

I. An Overview of Traditional Disability Benefit Schemes

For context, we summarize features of typical disability plans as they existed prior to 2000 and, in many countries, today. Most plans (except for some in Latin America), are managed by public agencies and financed on a pay-as-you-go basis out of payroll taxes, sometimes with supplements from the government's general budget. In a few cases mandated private plans also exist; they may be tied to employer, occupation or pension plan. Workers qualify for disability benefits when they are unable to carry out their regular work or other work or are able to do so only with great difficulty or pain, due to a medical impairment. Applicants present their cases, generally with evidence from their own personal doctors, and sympathetic agencies are often reluctant to deny claims. Some systems permit temporary and/or partial benefits but permanent benefits have been common and reassessments rare. If the individual returns to work benefits cease or are greatly reduced—so once he qualifies for a permanent benefit he almost never returns to work. In recent years these procedures have been challenged and, consequently, have begun to change. (See Table 1 for characteristics of traditional schemes and recent reforms).

Defined benefits

Disability programs pay defined benefits, usually at least 70% and sometimes 90% of recent wages, even in many countries where the old age system is defined contribution. Child allowances and in-kind benefits like housing and health insurance are frequently added (see Table 1). Essentially, if a person is “tagged” as disabled, this entitles him to benefits that are usually more generous than other non-workers such as unemployed individuals or early retirees. The presumption is that disability is an involuntary exogenous condition while unemployment and early retirement are more discretionary. Moreover, the disabled, being unable to work, are considered less subject to moral hazard and labor disincentives than are able-bodied individuals, so generous compensation supposedly does not have large distortionary effects.³ However, as we shall see, empirical evidence indicates that higher benefits increase the probability of moral hazard—they lead to greater reciprocity rates, less labor supply.

Spending, reciprocity rates and PAYG finance

Public plus mandatory private spending on disability benefits in OECD countries ranges from less than 1% to almost 6% of GDP, with an average of 2.6% of GDP, almost 5% of wages

(Table 3 and Figure 1). OECD countries, on average, spend one-third as much on incapacity as on old age pensions—but this too varies widely, from 10% to 90% (Figure 2). Public spending on disability is 10% of all social spending. Benefit recipiency rates vary from less than 1% to over 9% of the working age population, with an OECD average of almost 6% (Figure 3). Larger plans are found in middle- and high-income countries, especially Sweden and Norway, smallest in low-income countries such as Mexico and Turkey (OECD Social Expenditure Database—SOCX 2008).

This pattern is explained, in part, by the greater social security coverage in high income countries. Viewing the data as a percentage of old age spending helps control for coverage effects but large differences still remain. Wealthier countries may have a higher collective demand for disability insurance because they can afford to pay more for income stability and security. In poor countries there may be little distinction between those who are unemployed due to disability versus those who are simply unable to find jobs because of chronic unemployment and underemployment. Poor health conditions and education means that much of their populations are, in some sense, disabled. However, much of the variation is independent of national income and stems from differing design features among countries. Often the long run cost implications of these design features are not taken into account when they are adopted. This is partly because pay-as-you-go (PAYG) finance is commonly used.

Distribution of disability benefits

Although higher national wealth seems to increase the collective demand for income security, the main beneficiaries within a country are those with low permanent earnings capacities. People who self-report themselves as having a chronic health problem that limits their daily activities are likely to be older than the average population member, have less education and are not employed or work only part-time while employed. For these reasons, their income is lower and poverty rate higher than for the population at large. This is even more true of disability beneficiaries, but the benefit offsets the lack of wages and helps to forestall poverty.

In the U.S., disability rates are 3.5 times higher for those without college education than for those with (Chandra and Samwick 2008). In most OECD countries, the proportion without a full secondary education is double that for the disabled compared with the non-disabled population and this gap has been growing for younger cohorts (OECD 2010). Individuals with less education are generally in worse health due to life style factors and their jobs are physically

more demanding, less interesting. Moreover, the progressive disability benefit replaces a higher proportion of their wage than for those with more education; hence disability reciprocity status is more attractive to them and they are more likely to apply.

Contributions (or premiums) for mandatory disability insurance are usually not differentiated by risk category. Rather, they tend to be a fixed proportion of income, so low earners pay less than high earners. Because they have a greater risk of becoming disabled and receive a more generous replacement rate but pay a lower premium, less educated, low earning groups get a higher net expected value and rate of return from disability insurance. This helps to offset the regressivity of old age systems stemming from the greater longevity of high earners and makes the over-all social security system more progressive. The collective decision, apparently, is to insure against disability and also insure against the possibility of higher premiums if the probability of disability rises for an individual. This might be the choice about social policy that many individuals would make from behind the “veil of ignorance”—that is, before they have any information about their genes, life style and socio-economic status, all of which differentiate their probability of becoming disabled.

Age is also strongly associated with disability and disability benefit reciprocity. Older workers are more prone to chronic diseases and have accumulated more health impairments, which may impede their ability to work. Beyond that, disability systems often offer more generous treatment to older workers, giving them top-ups to benefits, which makes them more likely to apply (e.g. Spain). Gatekeepers sometimes apply a looser definition of relevant job or take into account labor market conditions for older workers with disabilities, making it easier for them to qualify (e.g. Luxembourg and Austria). As early retirement schemes are tightened, many older workers who previously would have retired early now take disability benefits instead. Disability costs will grow as populations age, especially if these policy biases remain.

Diagnoses

Originally disability insurance was intended for those who were suffering from well-defined debilitating illness such as cardiac problems and expected to die soon, but over time ambiguous chronic non-fatal conditions such as arthritis, pain and depression became eligible for benefits—and their prevalence ballooned. Currently, muscular-skeletal and psychiatric conditions constitute over half of all new entrants. The reasons for this shift are not totally known, but deliberate policy change is clearly part of the reason. For example, in the U.S., before

1984 to qualify for disability benefits a person had to have a listed medical impairment or other verifiable diagnosis, but after 1984 gatekeepers had to consider an applicant's "ability to function in a work-like setting," to place heavy weight on reported pain, to relax screening of mental illness, and to place major weight on information from the applicant's own doctor.⁴ These diagnoses are especially difficult to identify and measure objectively and therefore especially subject to moral hazard. The greater the probable benefit the greater the growth in these subjective categories, which helps to account for the high rate of reciprocity despite the over-all health improvement in the population (see below).

While the risk of disability is still strongly and positively correlated with age, more recently the inflow of young disabled has been growing most rapidly, since that is where the ballooning psychiatric problems are concentrated. It is unclear whether these conditions are temporary or permanent. If these individuals remain on benefits and out of the labor force for the rest of their lives, the cost to the disability system and the broader economy will be huge

Relationship to unemployment rate and other social insurance programs

Within a given country the inflow of disability recipients rises during periods of high unemployment. When job opportunities fall, employers are less likely to hire those in ill health. From the workers' viewpoint, disability benefits appear large and permanent relative to temporary and smaller unemployment benefits. This underlines the complex inter-relationship between disability insurance and other social insurance programs like unemployment benefits, early retirement and social assistance. Typically the same people are eligible for more than one of these programs and they apply for disability benefits if these are more generous. But when access is cut off to the preferred program, perhaps for fiscal reasons, rejected applicants flow to the others. Thus, changes in one part of the social insurance system affect the other parts, a fact that policy-makers must bear in mind as they make reforms.

II. Problems—

1) Disability is non-observable and subjective, hence classification errors

Disability benefits are intended for cases where health impairments severely limit work capacity. However, it is difficult to observe and objectively measure this condition, to define how severe the limitation is or how much difficulty the individual faces when working. At the

extremes, it is clear when a person is totally unable versus fully able to work and care for himself or herself, but in-between these extremes is a large grey area. A person who loves his or her job may find ways to continue despite a medical impairment, whereas a person who strongly dislikes it may find it very uncomfortable to carry on in the presence of health problems. (This may help explain why the disability benefit reciprocity rate is higher among those with less education, who have less interesting and physically more demanding jobs). Psychiatric problems and muscular-skeletal conditions like arthritis and pain, which currently predominate among recipients, are especially non-observable. An MRI may show lumbar stenosis but it does not show the pain which may or may not be emanating from that stenosis. Is a depressed person unable to perform on the job or, if encouraged to work, will this mitigate the depression? Within this grey region, subjective evaluations play a large role and differ among applicants and examiners.

Information available to gatekeepers is often limited—traditionally it comes from the applicant’s own doctor, who may not be fully educated about the connections between physical impairment and ability to work. A common reform measure requires an examination by the social insurance system doctors and/or vocational experts, instead of or in addition to the applicant’s own GP but this, too, does not give complete information. Because the difficulty in assessing the extent of work incapacity is at the heart of the implementation problems that disability systems face, we start by summarizing the evidence on high rates of inclusion and exclusion errors.

Classification errors based on audits

One method for measuring classification errors is to send in a team of experts to “redo” the initial decisions. This kind of audit is rarely done, no doubt because of its high expense. The classic study of classification error was by Nagi (1969), who conducted an intensive analysis of this issue for the U.S. Social Security Administration (SSA) in 1969. Nagi’s team (which included medical, social and vocational experts) evaluated 2454 recent disability applicants—some had been accepted and others denied. Teams of 5 experts visited homes, interviewed applicants, came to a joint decision on their disability status and compared these with SSA’s prior decisions. Assuming that the Nagi team was “right,” 48% of SSA’s denials were false while 19% of SSA’s awards were false. Over-all, 30% of SSA’s decisions clashed with the Nagi teams’ assessments. (Benitez-Silva et al 2004b). The evaluation of individuals with multiple or psychological impairments—a growing category—proved especially difficult. Few of the

applicants, whether accepted or denied, held jobs. Many denied claimants received some kind of public assistance in lieu of disability benefits, while some reapplied and later received benefits from SSA.

Another study for SSA found that only 36% of denied applicants had earnings ten years later, these earnings were well below the median and they had high mortality rates—again suggesting many false negatives. Still another study found 20% false negatives and 20% false positives. (For summaries of these studies and other evidence see Bound 1989 and Parsons 1991). Audits in other countries have also found errors of inclusion and exclusion in their welfare programs, although more tilted toward the former.⁵

Different approval rates among examiners, administrative judges and localities

A second method for measuring classification errors controls econometrically for different approval rates among various examiners in a system and exploits the fact that some examiners are systematically more lenient than others. This method requires detailed information about individual cases, including the identity of the examiner who evaluated the evidence and made the decision. A recent study of this sort in the U.S. found large variations in approval rates across examiners, after controlling for applicant characteristics (Maestas, Mullen and Strand 2010). Another study found substantial variations in the approval propensity of administrative law judges who hear the second stages of the appeals process in the U.S. (French and Song 2009). In both cases, this variation indicates substantial errors of inclusion or exclusion, although we don't know which since we don't know which decision was "correct." (These studies used this variation to estimate large labor supply responses among otherwise similar individuals who received or did not receive benefits simply due to the propensities of the examiner or judge to whom they were randomly assigned.)

Consistent with these results, we know that 50% of appealed denials are accepted at the first appeal stage, 75% of the remaining denials at the second appeal stage and an additional 30% of remaining denials at the third stage--further evidence of the ambiguity and subjectivity of disability determination in the U.S. The assessment changes as new examiners look at the data (Benitez-Silva et al 1999).

Differences in approval rates across states in the U.S., cantons in Switzerland, regions in Sweden and provinces in Chile that cannot be accounted for by regional characteristics further exemplify the subjectivity of the decision process and the existence of classification errors. In

Denmark, the ratio of minimum to maximum benefit award rates across municipalities is 1/3 (OECD 2008). Given Denmark's homogeneous health conditions, these divergent regional assessments of impairments cannot be correct; both false positives and negatives are likely.

System decisions vs. self-reported health status

A third way to measure errors of classification is to compare decisions of the system gatekeepers with the individuals' self-reported health status in household surveys like HRS. In HRS individuals are asked whether they have a medical impairment that reduces their ability to carry out the activities of daily life such as work. About 12% of the working age population self-report some incapacity in the U.S. and the numbers range between 10% and 20% in most European countries. This may be compared with benefit reciprocity rates of 6% of the working age population in the U.S. and 5-10% in most of Europe in 2008. In 26 OECD countries as a group, around 2005, three quarters of people who reported themselves as disabled had no disability benefits, one quarter got no public benefit of any sort and 9% had no benefits or work (possible errors of exclusion). In roughly the same time period, about one-third of all disability benefit recipients did not self-report themselves as disabled (possible errors of inclusion). This disparity between reported disability and benefit reciprocity may be due to differing definitions by the individual and the system, to the possibility that the disability is temporary or partial while benefits are only for permanent or full cases, to the absence of reassessments or to the fact that disability insurance covers only part of the population. The excluded group is virtually non-existent in countries like Norway and Sweden, where eligibility for benefits does not depend on contributions and former employment status. (Data are from OECD 2010 and 2003; for more country details see OECD 2006, 2007, 2008 and 2009).

Benitez-Silva and colleagues (2004a and b) compared disability awards in the U.S. between 1992-96 (after all appeals) with the applicants' self-reported disability status. They found that about 20% of SSA's acceptances are given to individuals who do not claim they are disabled in HRS, while 60% of denied applicants self-report they are disabled. The majority of these work, but for fewer hours and lower wages than those who do not self-report they are disabled. Several analysts conclude that, given the large proportion of apparent false negatives, screening in the U.S. may be too stringent and fails to give benefits to many who should get them. It might be preferable to give a smaller benefit to a larger number of applicants, holding the over-all budget constraint constant.

Of course, self-reported disability may be biased. Individuals who expect to apply for disability benefits are likely to report themselves as having health problems that make them unable to work even if, in fact, they can. They may not trust the supposed anonymity of the responses. Additionally, self-reported health status may not be comparable across individuals, socio-economic groups and countries that have similar objective health conditions.⁶ Countering these objections it is argued that individuals have little incentive to misreport in HRS, since these answers are used for research purposes, not for SSA determinations. It appears that self-reported disability is a powerful predictor of individuals' decisions to apply and appeal, as well as SSA's ultimate decision to award benefits and are also correlated with clinical measures (Benitez-Silva et al 2004a and b, Bound and Burkhauser 1999, Low and Pistaferri 2010).⁷ Nevertheless, for the above reasons they should be used with caution as a measure of "true" disability and classification errors.

System decisions vs. mortality rates and body mass

A fourth approach to errors of inclusion and exclusion is to compare gatekeeper decisions with other evidence on health status. In a cross-sectional study of disability decisions at the state level in the U.S., denials were targeted toward the more healthy applicants, as proxied by their body mass index (Gruber and Kubic 1997). In another inter-state study, subsequent mortality rates of those who were granted benefits exceeded the mortality rates of denied applicants, and mortality among those who appealed their denials was greater than among those who didn't appeal (Parsons 1991b). Further along these lines, in the new system adopted by Chile in 1981, which gave a larger role to private pension funds and insurance companies in the assessment process, successful claimants were found to be less healthy (as proxied by their subsequent mortality rates) than the over-all population in their age group, while this was not true of the previous public system (James, Edwards and Iglesias 2009). The over-all conclusion from these studies: careful screening can distinguish between the more and less healthy on average, but many disparities remain between benefit awards and objective measures of health status.

The trade-off between errors of inclusion and exclusion

More stringent screening could reduce false positives but might increase false negatives; policy-makers, therefore, face a trade-off between errors of exclusion versus inclusion. Within a given budget constraint, should they err on the side of rigorous screening, which allows a larger benefit to be paid to those whose disability is more certain, or should they pay a smaller benefit

to a larger number to avoid false negatives? More careful screening that tries to cut down on both errors implies higher administrative costs. At some threshold, it is not worth incurring these costs. These difficult choices would not arise if the state of being disabled could be readily observed and measured. As discussed in the following sections, recent reforms have tried to reduce errors of inclusion and exclusion by developing new sources of information, including self-selection mechanisms, monitoring by employers and job counselors, participation by pension funds and insurance companies and by extending the period of observation through periodic reassessments—all of which involve higher administrative costs.

Low-income countries face a scarcity of well-trained examiners and administrators needed for accurate assessment. Their error rates may be still higher than those mentioned above. The complexity of running disability programs and defining who is disabled suggests that the disability programs should be modest and narrowly targeted in these countries. (For a discussion of administrative problems see Goldblatt 2009 and Rao 2009).

III. Problems—2) Moral hazard, agency problems and social norms

Moral hazard

Economic incentives affect the response to insurance. When people have insurance, their potential loss falls, so they are more likely to undertake risky behavior—a phenomenon known as “moral hazard.” So, if individuals are insured against loss of income due to disability, we would expect they are more likely to stop working once they qualify for the benefit. In contrast, in the absence of insurance they are more likely to continue working even if this causes them pain and disutility, because they need the income. Moreover, a higher degree of insurance (larger benefits, easier eligibility) implies a higher return on the application investment (time, money, lost earnings) and therefore an augmented likelihood that they will choose benefits over work.

Several empirical observations are consistent with the existence of moral hazard as a consequence of disability insurance:

1. Labor supply falls and disability reciprocity rates rise when benefits increase.
2. Low earners, who get a higher replacement rate, are more likely to file disability claims than high earners, who get a lower replacement rate.
3. Disability claimants rise during periods and in regions with high unemployment rates.

4. An increase in the incidence of hard-to-diagnose medical conditions occurs when benefits increase and denials decrease.

5. Practically no one on permanent rolls gives up the benefit to return to work.

Parts IV and V summarize the evidence on each of these points.

These empirical observations have several possible causes, including the fact that many disabled persons find it exceedingly difficult to work, especially difficult to find jobs during periods of high unemployment, and practically impossible to return to the labor force after a long absence. Nevertheless, each of these observations is also consistent with moral hazard: workers respond to system incentives. They are more likely to label themselves as “disabled” and unable to work when they have less to lose by doing so, due to insurance. Bound et al (2004) estimates that the total cost of providing another \$1 of disability benefits is \$1.50, because of moral hazard.

Agency problems

Closely allied to moral hazard are agency problems: other parties use the disability insurance system for their own interest, which sometimes conflicts with the system’s interest. For example:

1. Doctors and other gatekeepers may apply a lax definition of disability in order to keep their clients happy. If a GP does not give his patient a letter attesting to disability, the patient may find another doctor who will oblige.

2. Employers who want to get rid of redundant or low-productivity workers may help them apply for disability benefits. The worker gets more leisure with little loss of income, and the employer frees up a position without firing hassles or severance pay.

3. Insurance companies may help clients to get on the public rolls, to avoid responsibility for paying privately.

3. Politicians who want credit for a low unemployment rate may make it attractive for workers to sign up for disability instead of unemployment benefits.

4. Union leaders find disability programs a convenient pathway to early retirement for their members.

5. State and local governments shift residents to disability programs (which are often nationally funded) to reduce their social assistance obligations (which are usually locally funded). (For evidence from the U.S. see Rupp and Stapleton 1995).

The conflicts faced by Swedish social insurance officers who are torn between demands of the workers and rules of the system, and who find it difficult to evaluate medical diagnoses, are documented in Soderberg and Mussener 2008. While gatekeepers have the responsibility to identify successful disability applicants, governments need to monitor the gatekeepers to make sure they are putting forth effort to avoid errors and not imposing their own preferences, which may differ from the system's preferences. The benefits from tagging beneficiaries must be weighed against the costs of monitoring the gatekeepers. Allowing beneficiaries to work without penalty and observing how many do so may be an inexpensive way to pick out those gatekeepers who are being too lax or too stringent (Boadway, Marceau and Sato 1997)—but if the gatekeepers know this device is being used they may discourage such work.

The role of social norms and contagion

In the absence of objective measurements, social norms play an important role in defining the level of incapacity that should qualify workers for disability benefits. Once a sizeable group adopts a lax interpretation of incapacity, this definition is contagious and spreads to others—which seems to have happened in countries with high cost and recipiency rates. In a Swedish test of this hypothesis, one district in a city was exposed to less monitoring, which increased its residents' sickness absences from work. But members of the non-treated group in the same city also increased their work absences, albeit not as much. In contrast, absences for groups outside the metropolitan area did not change. Contagion also applied to workers who came from the same immigrant community as those in the treatment group—they, too, increased their absences from work. So new social norms are established that redefine what constitutes an incapacitating illness and when it is appropriate to take sickness and disability leave, given the absence of objective measures (Hesselius, Johansson and Vikstrom 2008).

Implications for policy design

If disability is hard to measure objectively and moral hazard problems are common, then labeling a subset of jobless individuals as disabled rather than unemployed or on social assistance and giving them higher benefits may have several perverse effects. It may encourage them to stop work to qualify for the disability tag, differentiate benefits across individuals who are really quite similar, lead other interested parties to divert beneficiaries to disability and raise total costs. Recent structural reforms have been designed to reduce the disparity between disability and unemployment benefits, improve information flows and implement incentives that

better align the interests of the disability system and key decision-makers. This often means greater reliance on employers, pension funds and insurance companies in the private sector.

IV. Problems—3) High costs and reciprocity rates: sensitivity to benefit size and screening

Most disability insurance is publicly managed and financed, but in some countries employers are required to provide sickness leave and incapacity. We include public and mandatory private spending in this discussion. Design features of disability systems, especially generous benefits and lax screening, have led to high costs and reciprocity rates in many countries. Policy-makers are beginning to take these effects into account.⁸

Costs of disability insurance

Costs of disability insurance are less of 1% of GDP in low-income countries that have small systems with only a small proportion of their population covered. They rise to 2-6% of GDP in high and middle-income countries. On average, 6% of the working age population (WAP) receive disability benefits in OECD countries. The Nordic countries and the Netherlands are on top of the spending and reciprocity list. In these countries—which are the healthiest in the world according to life expectancy and other measures—public plus mandatory private spending on disability is 60-90% as much as spending on old age pensions, one out of every 8-10 working age persons receive disability benefits and over a third of all men retire on disability benefits before reaching the normal retirement age (Table 3 and Figures 1-3; data from OECD 2010 and OECD Social Expenditure Database 2008; also see OECD 2006, 2007 and 2008 and Haveman and Wolfe 2000). Yet, equally high-income countries like the U.S. and Japan have much lower disability rates, as do some middle-income countries in Latin America, like Chile. In most cases costs and reciprocity have been rising but some countries, such as Poland and the Netherlands, have seen dramatic cuts over the past decade.

Why the divergence across countries and changes over time?

Large health improvements should imply less disability, but instead we seem to have more disability and countries with similar health indicators have vastly different benefit reciprocity rates. What explains the huge divergence across countries and the sharp changes over time? Population aging may be part of the answer, since disability tends to be concentrated in older age groups—but in fact it is only a small part and the growth in disability rates has been

most rapid in younger age groups. Old age pension systems have tightened their early retirement requirements, which may have shunted more people into disability benefits—but that, too, is a very partial explanation since many workers continue to retire early in countries with high disability rates.

In fact, studies have shown that variations across countries and changes over time are driven by disability policies, especially policies regarding benefit generosity and stringency of screening. For example, Borsch-Supan (2010) regresses the probability of receiving disability benefits on country-specific policy variables like country coverage rates, minimum disability level required for benefits, benefit generosity, type of medical and vocational assessment required, as well as a variety of individual-specific demographic and health variables. While the latter are significant, most differences across and within OECD countries are due to differing disability policies. The level of a country's unemployment benefits also strongly affects the probability of being on disability rolls. Normalizing for demographic and health variables hardly changes the cross-country relative picture, but normalizing for policy characteristics argely equalizes predicted enrollments.

Disability applications respond to benefit levels and replacement rates

Numerous studies estimate the elasticity of disability applications with respect to benefit levels in the U.S. Most estimates lie between 30-70%, while the elasticity of awards is 30-40% (Bound and Burkhauser 1999). This sensitivity to benefit generosity is much higher for the moderately disabled than the severely disabled, who can't work and don't have any other income (Low and Pistaferri 2010). The increased value of Medicare and Medicaid, for which disabled beneficiaries are eligible, also induced more applications in the 1980's and 1990's (Stapleton and Dietrich 1995 and Stapleton et al 1998). Autor and Duggan (2003) find a close link between disability enrollments on the one hand and regional variations in wage levels and disability replacement rates on the other hand. Applicants in low-wage states get a higher replacement rate from the system's progressive formula, are more likely to apply and the system's gate-keepers let them through. The fact that applications vary in response to benefit amounts is evidence of moral hazard--claiming the status of "disabled" is not exogenous but depends on rewards from the insurance system. At the same time, the fact that these applications are not weeded out by the screening process is evidence of the difficulty in objectively measuring the disability condition.

A high elasticity of reciprocity with respect to benefit size has also been found in the Netherlands and Germany (Aarts and de Jong 1992, Riphahn 1995).

A study by Campolieti (2002) based on the Canadian Pension Plan (CPP) and Quebec Pension Plan (QPP) throws light on the mechanism that produces these results. Using aggregate data over a long period of time, Campolieti shows that among beneficiaries, the incidence of hard-to-diagnose medical conditions such as muscular-skeletal problems increase substantially during periods when the replacement rate increases. Generous benefits have apparently induced people with ambiguous medical problems to apply and their ambiguity makes it difficult to screen them out. The benefit breakdown among various diagnoses appears to be endogenous to the benefits offered—evidence of moral hazard in the application and evaluation system.⁹ This may help explain the association between generous benefits, rising reciprocity rates and an increasing dominance of muscular-skeletal and psychiatric conditions in the disability systems of many countries.

Evidence from Sweden shows that benefit amount affects the probability and duration of sick leave (a common predecessor to disability benefits), even though this is supposed to depend strictly on medical condition. In 1991 Sweden reduced its replacement rate for sickness absence from 90% to 65% for the first 3 days, 80% between days 4 and 90, and reverted to 90% after day 90. As expected, the hazard of absence (probability of starting a sick leave for the “at risk” group) fell sharply. And the duration of absence spells in days 1-90 grew shorter due to the reduced insurance benefit. However, the reversion to the high replacement rate at day 90 had a perverse effect: the duration of absence spells grew longer for workers who were already on sick leave at that point. They continued their leave, with practically full pay, rather than returning to work and risk the possibility of getting sick again and re-starting leave with a lower replacement rate. Ironically, the over-all probability of being absent from work grew, as a result of this response to the changed benefit formula (Johansson and Palme 2005). In 1998 Sweden allowed collective agreements to augment the replacement rate after day 90 by 10 percentage points, financed by the employer. This provision applied to municipal workers but not to central government workers, thereby constituting a quasi-experiment. Municipal workers who had over 90 days of absence increased their sick leave by 4.7 days, thereby increasing the cost of the publicly provided sickness insurance by 3% (Hasselius and Persson 2007).

In addition to its impact on reciprocity rates, benefit size also influences the length of time that individuals with a moderate work-limiting impairment typically wait, before applying for disability benefits. They are likely to apply sooner if they have higher expected benefits; in the U.S., a 20% increase in expected benefits reduces the time to application by 1.2 years for men, 1.7 years for women (Burkhauser, Butler and Weathers 2001-02).

Applications and enrollments respond to screening procedures

Several studies have found sensitivity to tightness of the screening procedures. Years of work needed for eligibility, waiting period before benefits can begin, degree of capacity loss required and qualifications of assessors are all important design features, but the indicator used most frequently in studies is the rate at which applicants are denied versus accepted for benefit receipt. Higher denial rates can reduce reciprocity rates and costs through two paths: first, through constraints—some persons who want to get disability benefits can't and second, through incentives for behavioral change—a rise in denial rates implies a fall in expected benefits, so some individuals won't apply, since the very act of applying costs time and money.

Using variations across U.S. states in denial rates, Parsons (1991a) found a strong response in number of disability applications: a 10% increase in initial denial rates decreased applications by 4.5%. Self-selection seems to work: on average, those who don't apply are more able than those who do apply. A higher denial rate therefore permits larger benefits to be paid to the more disabled within the same budget constraint, as a result of self-selection. However, later studies that controlled for state unemployment rate and population aging cut the denial effect in half (Rupp and Stapleton 1995, Stapleton et al 1998). A 10% increase in state acceptance rates also reduces the time from onset of health problems to application for disability benefit by .5 years for men, .8 for women (Burkhauer, Butler and Weathers 2001/02).

When Austria tightened its eligibility criteria for older men, their disability enrollments fell by 5.5 percentage points and their employment increased by about 2 percentage points. At the same time, through spillover effects, unemployment benefit reciprocity increased by 3 percentage points and sickness insurance benefits by .6 percentage points—offsetting 2/3 of this decline (Staubli 2009).¹⁰ In a controlled experiment in the Netherlands, applicants in 2 out of 26 Dutch regions were subject to intensified screening, including direct observation by case workers. Disability application and reciprocity rates went down and this was not offset by a rise in unemployment benefits (de Jong, Lindeboom and van der Klaauw 2006). However, it should

be noted that this experiment took place in the context of far-reaching reforms in Dutch system, described in Parts VII and VIII. The effect of intensified screening might have been quite different if this were the only change.

Disability recipiency rate varies with unemployment rate

If the benefit recipiency rate varies with economic incentives, it should also vary with the unemployment rate in the broader economy. Employers are less likely to retain disabled workers during periods of high unemployment, if disability reduces productivity but wages cannot be adjusted accordingly due to legal or social constraints. New jobs are less likely to be offered to those in ill health—since healthy productive workers are readily available at the same wage rate. Disabled workers who are laid off will have a harder time finding new employment and any job that they do find may pay lower wages. Then, disability benefits provide higher replacement rates than they did before, if benefit formulae are progressive. Disability benefits are typically higher and longer lasting than unemployment benefits. Therefore, more individuals have an incentive to apply for disability benefits during periods of high unemployment. Also, employers who wish to lay off workers without imposing economic hardship or unrest may help them apply for disability benefits.

In Sweden in 1990-91, localities with higher unemployment rates had a higher incidence of sickness absence, after controlling for other factors (Johansson and Palme 2002). Analyses of data from Australia in the late 1990's show that a high unemployment rate increases disability applications and grants, the longer the person has been unemployed the more likely he or she will end up on disability benefits, and the failure to do some work while collecting benefits early on makes it less likely that the person will ever exit from the disability rolls (Cai and Wilkins 2007).

According to a series of pooled, cross sectional and time series analyses for the period 1980-93 in the U.S., a 1 percentage point increase in the unemployment rate increases disability applications by 4% and final awards by 2%; it reduces approval rates by 2 percentage points (Rupp and Stapleton 1995, Stapleton and Dietrich 1995). The assessment process apparently weeds out some marginal applicants but approves others who were induced by job loss to apply. Moreover, once placed on the disability rolls during a downturn, individuals tend to stay on even after the economy turns up, due to deteriorating skills and work disincentives in the system.

Disability applications skyrocketed in the U.S. during the recession of 2008-10, given the loss of jobs and limited duration of unemployment benefits, and the cost is likely to continue long-term.

In Appalachia during the bust following the coal boom of the 1970's and 1980's, disability recipiency rose sharply as local earnings and employment fell, implying that higher replacement rates were provided by benefits. This effect was greatest for low-skilled men whose earnings fell the most. The (negative) elasticity of benefit payments with respect to local earnings was about 40% (Black and Sanders 2002).¹¹ In Chile, where disability insurance is run by private pension funds and insurance companies, the probability of receiving a benefit rises during periods of high unemployment (James, Edwards and Iglesias 2009). More broadly, Borsch-Supan 2010 finds an important role for unemployment rates in explaining varying disability beneficiary rates across countries and time. This suggests that rising replacement rates and falling work opportunities lead more people to apply for disability benefits.

Demand for disability benefits depends on access to other social insurance (and v.v.)

When a potential disability beneficiary is also eligible for some other form of social insurance such as unemployment benefits, early retirement pensions or social assistance, he or she will choose the benefit that is easiest to obtain, lasts the longest, and provides the highest replacement rate. This is often the disability system. In addition, some countries (Austria, Netherlands, Sweden) have explicitly used disability as an alternative to unemployment compensation or early retirement, especially for older workers. The demand for disability benefits depends in part on size of and access to these other transfer payments.

Borsch-Supan (2010) shows the sensitivity of disability enrollments to availability and relative attractiveness of unemployment benefits across OECD countries: smaller unemployment benefits raise enrollments in disability programs. In Sweden, disability rates have risen among young people, in part because they seem to be using it as a more generous substitute for unemployment benefits (Haveman and Wolfe 2000). In the U.S., when state and local social assistance programs were reduced in the 1980's and early 1990's, disability applications and awards rose by large amounts. As a cost-saving measure for their own budgets, several states launched vigorous efforts to help potential welfare recipients apply for awards from the federally-financed disability program and Medicaid (Rupp and Stapleton 1995).

A similar interaction applies between disability and early retirement pensions. Since 1983, normal retirement age has been rising in the U.S. from 65 to 67, gradually, for successive

cohorts. This reduces the pension paid to early retirees, which is based on an actuarially fair adjustment from the now-higher normal age. However, the size of disability benefits did not change so they became relatively more attractive. Disability enrollments have risen steadily for successive cohorts since 1983 (Duggan, Singleton and Song 2005). The raising of retirement age for women to equality with men in Australia similarly increased their disability enrollments (OECD 2007).

The relationship runs both ways: when special access to disability benefits was cut off to older workers in Austria, more than half of the excluded individuals ended up on unemployment benefits (Staubli 2009). When disability beneficiaries were reassessed and many were taken off the disability rolls in the Netherlands during 2005-9, some of them showed up on the unemployment and social assistance rolls. The tightening of access to sick leave in Poland in 1999 reduced the disability inflow but increased applicants for unemployment and early retirement benefits. Similarly, instituting tighter pre-conditions for disability in Luxembourg starting in 2001 reduced disability recipients but much of this was offset by increases in unemployment recipients (OECD 2007). When Sweden abolished the easier eligibility conditions for workers age 60-64, disability spending fell but sick leaves rose (Karlstrom, Palme and Svensson 2008). When Sweden took action to reduce sick leaves, unemployment benefits rose, and it rose most in regions where sick leaves fell most (OECD 2009). Some money might still be saved by these measures if disability benefits are larger than unemployment and social assistance benefits; but the long term savings are lower than initially appears.

Relationship to moral hazard and non-observability of disability

The elasticity of recipiency and costs with respect to relative benefit size, screening tightness and macro-economic conditions results from moral hazard combined with the ambiguous and subjective nature of disability. Claiming the status of “disabled” is not exogenous but depends on rewards from the insurance system—evidence of moral hazard. More claims pass through the screen in generous systems because of the difficulty in objectively assessing the disability condition. Recent reforms attempt to tackle this issue by tightening eligibility conditions, spending more on administration to get a more accurate decision, improving the information available to screeners, reassessing more frequently and changing incentives facing workers.

V. Problems—4) Low labor force participation rates of the disabled

In OECD countries, disabled persons have employment rates that are only 60% as high as the non-disabled and their unemployment rates are twice as high. When they work it is likely to be part-time and at low wages (OECD 2010). The picture is even worse for disability beneficiaries. On average, only one-third work and their incomes are less than the low threshold that is allowable for benefit recipients. This is, in large part, because a serious health problem impedes their work. But design features in the disability insurance system also discourage individuals from using their remaining work capacity. Incentives in the disability system not only have financial consequences, they also affect resource utilization and real output in the economy.

How does the disability system affect labor force participation (LFP)?

First of all, eligibility for a full disability benefit generally requires that the person is unable to work. So anyone who plans to apply must stop working (the planning effect). Presumably most claimants stop work involuntarily but those who could work, albeit with great effort, don't put forth that effort once they decide to apply for disability benefits. This effect is greatest among older individuals, who have a lifetime accumulation of disabling health conditions. Second, once the person qualifies for a benefit, he knows it will stop if he works beyond a threshold—that is, he faces a high (often close to 100%) implicit tax on work, which discourages him from trying (the incentive effect). Third, the benefit provides an income that substitutes for wages and can be spent, in part, on leisure (the income effect). The larger the potential benefit, the greater the incentive and income effects are. Fourth, return to work becomes more difficult as time out of work rises—skills depreciate, the work habit is lost and access to social networks that help to find employment deteriorates. As a result, once people qualify for disability benefits they tend to remain there. In most countries only 1-2% of all benefit recipients leave the disability rolls annually for reasons other than old age pensions or death. The disability system becomes a benefit trap; and older people are most likely to be caught in the trap. Since the prevalence of disability grows with age, the disability system has been given as a major reason for the falling participation rate among older individuals.

Stapleton and Burkhauser (2003) examine a number of alternative reasons for the increasing number of individuals who don't work due to disability in the U.S. over the last two decades—population aging, increasing job stress, employer unwillingness to hire the disabled

after the Americans with Disabilities Act (ADA) was passed, etc. They conclude that the major explanation is the expanded eligibility for and size of disability benefits—extensions for mental illness, heavy weight to pain in assessments and rising replacement rates for low earners. The new rules led to expanded numbers on disability rolls and once on the rolls they faced a strong incentive not to return to work. This explanation is consistent with Borsch-Supan's analysis of varying benefit recipiency rates in OECD countries.

Disability benefit amount and replacement rate affects LFP of older workers: U.S. studies

To quantify the responsiveness of labor force participation rates to benefit generosity, ideally one would randomly expose individuals to different benefit formulae and analyze their differential responses. But such random exposure to different public policies is rare. Usually, a single formula applies within a country. As a second best, analysts look for an appropriate control group whose members face different benefit options, and measure their differential responses while econometrically controlling their differences, especially differences in health status. Most of these studies use data from the U.S. and analyze how much of the declining labor force participation rate among older men can be attributed to the growth of disability recipients. Different results stem from the choice of control group, how health status is measured and how one separates out the impact of high replacement rate versus low earnings history and taste for work. We describe these studies in some detail because they illustrate the methodological challenges that make a definitive judgment difficult. In general, the studies agree that design of the disability system has a strong impact on labor force participation of older men, but the magnitudes of this effect vary widely.

One of the most widely cited studies (Parsons 1980 and 1982), uses panel data to compare participation rates among individuals facing high and low wage replacement rates from the disability system, with subsequent mortality rates proxying health status. He shows that the replacement rate has a powerful effect on labor supply—the elasticity of labor force non-participation with respect to the replacement rate is 63%. When the coefficients from his participation rate equation are applied to the demographic and benefit changes that took place during 1948-76, they explain virtually the entire fall in participation rates among older men over that period. The largest declines in participation rates are experienced by low earners, who got the highest and most sharply rising replacement rates from the disability system.

Bound (1989 and 1991) strongly criticizes Parsons' methodology and results. He replicates the Parsons study using the same data but compares labor force participation in the beneficiary group with that in a control group--applicants for disability who were denied benefits at the initial screening. He also changes the health measure, using self-reported health rather than subsequent mortality. Although this control group probably had more work capacity than disability recipients, the majority did not work after they were denied benefits and those who did work did so only part time, earning less than half the median male wage. They reported that poor health prevented them from working. By inference, this would have been even more true of disability recipients, so Bound argues that the replacement rate was only of secondary importance.

Bound proposes that the lower participation rates among individuals with higher replacement rates may simply be due to the fact that the U.S. benefit formula is progressive. Therefore low earners get higher replacement rates. And low earners may withdraw from the labor force for other reasons--because they have less pleasant jobs or a low taste for work which made them low earners in the first place. In other words, the individual's replacement rate stems from the same underlying forces (earnings history and taste for work) that determine his willingness to apply for disability benefits and drop out of the labor force. Thus the replacement rate is correlated with the drop-out rate but does not cause it. It is difficult to separate these effects in a progressive system. Bound's calculation of the elasticity of male labor force non-participation with respect to the disability replacement rate is less than half that of Parsons.

Parsons (1982 and 1991a), in turn, attacks the use of self-reported health as a control variable, arguing that it is endogenous and influenced by economic incentives. As evidence: the data show that people who face higher replacement rates and consequently plan to apply for disability benefit report themselves as being in poor health. Mortality rates are a more objective measure. He also attacks the use of rejected applicants as a reasonable control group, on grounds that the majority plan to appeal or reapply and therefore don't work as much as they would in the absence of a disability program. Other contributors to this stream of literature find elasticities of non-participation ranging between 20% and 80% (see Leonard 1979 and 1986, Slade 1984, Haveman and Wolfe 1984 and 2000, Chen and van der Klaauw 2008).

A more recent attempt (Maesta and Yin 2008) uses a different approach, based on the fact that the implicit tax on work by disability recipients disappears at age 65, when beneficiaries are

transferred to the old age system and are allowed to work without any penalty. The participation rate of former recipients relative to non-recipients rises by 10 percentage points at that point. Relative hours worked and annual earnings also rise. This probably underestimates the total labor supply impact of the disability system, since the responsiveness may be greater at younger ages, once recipients have been out of the labor force it is difficult for them to return, and the income effect from the pension remains for both groups.

Approaching this topic from yet another angle, Angrist, Chen and Frandsen (2009) study the reasons for the large growth in the veterans' disability compensation program in the U.S., as the number of Vietnamese veterans grew. Over-all, the health and work of Vietnamese veterans versus non-veterans are similar. However, among men with low earning potential, Vietnam veterans have much lower participation rates than non-veterans. Their explanation—veterans' disability benefits provided a very high replacement rate (sometimes 100%) to less-skilled men, and this led them to enroll in the disability program instead of work.

Disability benefit amount affects LFP: Canada

Gruber's (2000) study of the impact of disability insurance on labor supply in Canada probably uses the best treatment and control groups and therefore reaches the most widely accepted conclusion on this issue. In most countries all workers face the same benefit schedule, so it is impossible to compare the response of the same type worker (same earnings history and taste for work) to differing schedules. In contrast, Canada operates two separate disability programs—QPP in Quebec and CPP in the rest of Canada. Since the early 1970's, benefits grew more rapidly in QPP. But in 1987 CPP raised its replacement rate by 36%, to equalize benefits. Gruber studies the labor supply response by looking at the change in participation rates between these two regions after 1987. If disability recipients are truly disabled and unable to work, that is, if there were no moral hazard, there should be little or no response. In fact, he finds a large response for male workers age 45-59. Non-employment rose by 2.3-2.7 percentage points, which is 10-12% of their baseline non-employment rate. This means the elasticity of non-employment with respect to the replacement rate is about 30%. But another study showed that non-employment and non-participation did not change in 1973, when QPP increased its benefit size. This may be due to the fact that eligibility requirements and medical screening were much more stringent at that time, allowing less space for moral hazard. Thus, the impact of screening and benefit generosity may interact (Campolieti 2004).

Tightness of screening affects work probabilities

Previously we saw that tighter screening reduced the incidence of disability benefit recipients. It also reduces the probability that individuals will stop working. In an experiment in Sweden, the requirement of a doctor's certificate for sick leave was postponed from day 8 to day 15 in the treatment group. The duration of sickness absence increased from 11.9 to 12.6 days; the probability of returning to work was the largest one day before the deadline for the doctor's certificate. (Hesselius, Johansson and Larson 2005).

Gruber and Kubik (1997) studied the impact on participation rates of an increase in denial rates in the U.S. in the late 1970's. Due to a funding crisis, all states raised their rejection rates by 30% on average, but by varying amounts in different states, allowing an estimate of quantitative effects. Not surprisingly, Gruber and Kubik found that a rise in denial rates leads to higher participation rates—a 30% rise in denial rates produces a fall in labor force non-participation of 1.4 percentage points or 8% of the ex ante rate for men age 45-64. This implies an elasticity of non-participation with respect to denial rates of 28%. Denials were targeted toward more healthy individuals (with higher body mass index) and these were the ones whose labor force participation rates increased the most.

Analyzing the impact of varying denial rates among different disability examiners in the U.S., Maestas et al (2010) find that if the marginal person who was accepted instead had an examiner who denied his claim, his labor force participation rate would, on average, be 21 percentage points higher and he would earn \$1600-\$2600 more per year. This effect varies by medical condition and is much larger for those with less severe impairments, who have greater residual capacity to work. Applying a similar strategy to variations in denial rates by administrative law judges at the appeals stage, French and Song (2009) find that if a marginally accepted applicant had, instead, been assigned to a stricter judge who denied his claim, his labor force participation rate would have increased by 14 percentage points, on average.

Studies from Canada produce similar results. In 1989 a legal decision forced the Canadian Pension Plan (CPP) to permit the use of socioeconomic conditions (regional unemployment rate and worker's skills) in determining eligibility for CPP disability benefits. In contrast, eligibility criteria were unchanged in the Quebec Pension Plan (QPP). Campolieti finds a 1.5% reduction in the labor force participation of men age 45-64 in CPP relative to QPP after eligibility was loosened. This is attributed to an influx of beneficiaries with ambiguous medical

decrease benefit reciprocity rates and conditions who could have worked, albeit with difficulty. A steady increase of new beneficiaries who didn't work occurred in CPP but not in QPP, until CPP reformed its procedures in 1995 to disallow socio economic conditions as an eligibility criterion (Campolieti 2002 and 2003).

The impact of recent reforms

As discussed in detail in Parts VII and VIII, the basic thrust of recent reforms is to reduce the implicit tax on work, confront potential beneficiaries with a choice between different jobs rather than a labor-leisure choice and utilize pecuniary incentives from the private sector to dampen costs and labor force withdrawal. Disability reciprocity rates have responded to this approach. However, so far pro-work incentives have not been large enough to stimulate greater labor market participation by the disabled.

VI. Reforms—1) Changing the system's parameters

When a disability program is started or reformed, a number of design issues must be addressed that determine who is eligible to become a beneficiary, how large are benefits, who provides information and serves as gatekeeper, how stringent is the screening process, are periodic reassessments required and how does the disability system compare with other social insurance programs? We call these design features "system parameters." In this section we discuss how countries have been modifying these system parameters, in an effort order to avoid the problems just discussed—to increase accuracy while decreasing moral hazard and costs. Part VII deals with more basic changes that increase pro-work incentives for the partially disabled. Part VIII describes major structural reforms undertaken by several countries that have shifted informational and financing responsibilities to private sector parties (employers, pension funds, insurance companies) who have a direct pecuniary incentive to control reciprocity and costs. Table 1 summarizes key characteristics of the systems in countries discussed. Table 2 sets forth basic design questions and Table 3 lists outcomes that should be regularly monitored. Figures 1-5 display some of these outcomes—on spending, reciprocity rates, beneficiary inflows and outflows for a sample of countries.

Eligibility

Is eligibility for contributors only or also for non-contributors? Since disability programs are insurance schemes protecting against lost earnings from work, they are primarily relevant to individuals who are in the labor force and contribute to the system. When such individuals are unable to work because of a health impairment, they become eligible for benefits in lieu of wages. These are usually provided through the public sector, although mandatory private programs also exist. This is the core of most mandatory disability benefit schemes.

In addition, a few countries (Denmark, Norway, and Sweden) have flat benefits for the disabled, regardless of their work experience. Women are much more likely to be eligible for benefits in these systems, one reason for the high reciprocity rates in these countries. Most other high-income countries have small means-tested disability benefits for non-contributors, financed out of general revenues. The non-contributory programs are the 0 pillar of the disability system, aimed at poverty-prevention, while the contributory programs are Pillars 1 (public) or 2 (private), insuring against earnings loss.

In low and middle-income countries many workers are in the informal sector, where they do not contribute and are not eligible for benefits from the contributory scheme. This is a major reason for the low reciprocity rate in these countries. How long must an individual work in the formal labor force before she is covered by the contributory scheme? How long can she be out of the formal labor force before she loses eligibility? Easier eligibility—that is, quick coverage once contributions start and slow loss of eligibility after they end—means broader insurance but higher cost. It also leads to adverse selection, with workers trying to enter formal employment when their probability of applying for benefits is high but avoiding contributions when their probability is low. This raises the average annual price for everyone else and may eventually make the system financially non-sustainable.

Eligibility periods vary widely. In many European countries (Germany, Austria, Italy) five contributory years are required and in Latin America three years are typical, although these can be spread over a much longer period. In Chile the individual must have contributed within the past 12 months, and made at least 6 payments in the year preceding the last contribution. Chile avoids the adverse selection danger by counting as 0's all non-contributing years within the last ten, when calculating the reference wage on which the benefit is based. Thus, a person who has worked formally for only two years gets 20% of the benefit he would have gotten if he had contributed for ten years. This turns the on-off eligibility switch into a continuous function and

minimizes the effect of adverse selection on total spending but it also produces low benefits for many disabled individuals. These benefits are supplemented by the government's minimum pension guarantee, basic benefit or social assistance (James, Edwards and Iglesias 2009).

In countries with large informal sectors it seems wise to choose cautious rules for determining eligibility and reference wage that deter adverse selection for the contributory scheme, while leaving non-contributory arrangements to protect those who didn't contribute regularly but are most in danger of poverty.

Is eligibility based on the incapacity to do one's regular job or any job? At one extreme, benefits could be granted to people who are unable to continue doing their present job. A pianist who loses a finger would then become eligible for a disability income for life. In the past, this definition was common, but countries have been moving away from it in recent years to cut costs and encourage work by those who still have considerable capacity. At the opposite extreme, full benefits could be granted only to those who are unable to do any job. Countries have been moving in this direction for their disability programs, while retaining the own-job test for temporary sickness leaves.

For example, Norway abolished its own-occupation test in 1991, Germany in 2001. The definition in Korea, Turkey, Canada and the U.S. refers to any work. The UK applies an own-occupation test for the first 28 weeks of sickness leave but a broader definition thereafter. In Denmark and the Netherlands all jobs are considered, while the estimate of capacity loss also takes account of previous employment and education. In Poland's new system, total disability means that the individual can't perform any work while partial disability means he is unable to perform his usual work. But Mexico, Portugal and Spain still refer to the person's own-occupation as the criterion for disability (OECD 2003, 2010).

Broadening the definition of an acceptable job reduces the number of individuals who are eligible for benefits, reduces insurance costs and puts greater emphasis on work, for those who have the capacity to perform some job. When Austria shifted from the similar-occupation test to a reasonable-occupation test for older workers, disability enrollments declined by 5.5% and employment rose by 2% in the relevant age range (Staubli 2009). Given the cost-saving, this more limited view of eligibility is probably consistent with the mandatory insurance that most citizens would choose, in their dual identity as tax-payers and potential recipients. However, it may sometimes require individuals to perform jobs very different from and having a lower social

status than the jobs they trained for and expected to do. Persons who have invested in many years of advanced education for their jobs may wish to add supplementary insurance on a voluntary basis to cover their “own job” or their employers may finance this added protection.

Are the partially disabled eligible for benefits? Some countries (such as Canada, Mexico, Turkey, the UK and U.S.) grant only full benefits, although being “fully disabled” is defined differently in each case. The presumption is that these beneficiaries are substantially unable to work—an assumption that is at odds with the effort in some of these same countries to get them back to work. Many countries (e.g. Switzerland and the Nordic countries) grant partial benefits for partial disability. Reciprocity rates tend to be higher in these countries (OECD 2003 and 2010). The income of partial beneficiaries may be low, yet they are often penalized by benefit loss if they try to increase it by work. A few countries (Australia, Luxembourg) have recently tightened the definition of full disability while those with partial work capacity are eligible for job search assistance and unemployment benefits if they don’t find a job. One problem is that their partial loss of capacity may be long term, unlike unemployment status, which is presumed to be short term. Perhaps the most interesting reforms try to separate out the (almost) fully disabled from the partially disabled, granting full benefits to the former and long-term wage supplements to the latter. These schemes, as typified by the Netherlands and Denmark, are discussed more fully in Part VII.

Benefit amount and replacement rate

An extensive literature summarized in Parts IV and V documents the sensitivity of reciprocity rates, costs and labor force participation to benefit generosity. A higher wage replacement rate for the disabled implies more insurance and less risk for the individual but higher costs for the group as a whole. How much insurance would rational citizens wish to mandate before they know their individual risk profiles, given that benefits received by those who turn out to be disabled will be financed by premiums paid by the able and may have distortionary effects?

As a minimum, even among the less risk averse, most citizens would probably favor a level of insurance that keeps them out of poverty. This implies that low earners will get a relatively high replacement rate as, indeed, we find in most countries. Citizens who are more risk-averse might prefer a higher level of insurance that guarantees a continuation of their customary standard of living, maintenance of their social status and a reasonable return on their

education. However, this is much more expensive, probably more than the average citizen would wish to support. It is also regressive, since high earners generally have higher social status and education.

The prospect of moral hazard and agency problems further limits the optimal degree of insurance. The possibility that some individuals might find ways to game the system—qualifying for benefits, withdrawing from the labor force, consuming more than they produce and thereby raising the costs that non-beneficiaries must finance—reduces the benefit size favored by the average citizen who does not expect to game the system. The larger the probability of moral hazard the lower the optimal benefit size, below the level that would be optimal with full observability of disability and without behavioral adaptations to insurance.

Initially replacement rates in the contributory schemes of many countries were very high—consistent with the belief that moral hazard was a small concern. But accumulating evidence regarding the responsiveness of claim applications and labor force participation to benefit size suggests that moral hazard is larger and optimal benefit size smaller than initially thought (see above). Currently the average benefit is 50-60% of the average wage (net of taxes) in the contributory schemes of most OECD countries. But this average obscures broad variations. Replacement rates are 70% for the fully disabled in the Netherlands, Switzerland and Scandinavia, and even higher when widespread occupational top-ups are included. Among low earners, they usually exceed 80% even sometimes 100%, given the progressiveness of the benefit formula and the availability of means-tested housing allowances, child supplements and social assistance. Benefits are lower (but replacement rates still high for low earners) for the partially disabled, in flat benefit systems and in non-contributory schemes (OECD 2010).

Countries have been reluctant to cut replacement rates despite moral hazard because this reduces the insurance value for those who are truly unable to work. As discussed in Part VII, the most promising recent reforms distinguish between the majority who retain some earning capacity versus the minority who don't, retaining a generous benefit for the latter and partial benefits for the former, who are also allowed to keep any wage that they earn.

Screening

How stringent is the screening mechanism? When a person applies for disability benefits, this requires an assessment to determine whether he has a health problem that reduces his earning capacity and, if so, by how much. What should this procedure consist of and how

stringent should it be? Loose standards produce many false positives, thereby raising costs or allowing only small benefits per person. Some countries (e.g. Sweden and Austria) applied laxer standards for older workers, increasing their reciprocity, but this option has been closed in recent years. Tight standards reduce inclusion errors and allow higher benefits within a given budget constraint, but may incorrectly exclude many individuals who have limited work capacity. Moreover, stringent procedures incur high administrative expenses and appeals. For example, in the U.S, which has a complex multi-level assessment procedure, administrative costs were 2.7% of disability benefits, compared with .8% of old age benefits in the 1990's (Parsons 1996).

The most common indicator of stringency is the denial rate. Among applicants in OECD countries, the rejection rate in 2008 varied from only 18% in Norway to 65% in the U.S., with the median around 45%. Claims that were denied at the initial stage can be appealed but usually approved claims cannot be appealed—creating an inclusion bias. (Chile allows approved claims to be appealed, thereby eliminating this bias and reducing disability inflows and recipients). Successful appeals among the rejected were generally quite low—below 15% in OECD countries—with a few exceptions such as Denmark and the U.S. (OECD 2003 and 2010). Of course, this rate depends on quality of applicants as well as stringency of screening, both of which vary across countries. Low rejection rates in Norway are one reason for its high inflow rates while rising rejections in Poland are one reason for the sharp decline in disability inflows after 1999; but this has also led to many appeals and early retirements (OECD 2006).

Rather than stringency, it seems efficient to emphasize accuracy—increasing unbiased information and/or removing biased information. This could reduce false positives and its attendant costs without increasing false negatives. Reforms have been moving in this direction.¹²

Source of medical information—individual's own GP or system's physicians? In the past it was common to base the assessment on information from the individual's own GP, but recently systems have been adding checks by the system's physicians, as a more unbiased source. Personal GPs are likely to be sensitive to their client's wishes and unfamiliar with the interaction between medical conditions and vocational capacity. An applicant can shop around for a sympathetic GP. Instead, the medical assessment is now generally carried out by one or more physicians hired by the social insurance system in place of or in addition to the individual's own GP. The Netherlands and Germany have shifted to primary reliance on insurance system physicians—as a way to get greater accuracy and control costs at the same time (OECD 2003

and 2008). Teams of insurance experts carry out the medical assessment in Spain, Austria and Korea. As part of its major reforms in 1999, Poland imposed stricter reporting requirements for GPs and reassessment by system doctors in case of prolonged sickness absence. For this and other reasons, absences fell sharply in 2000 and continued to fall for the next few years. This in turn led to a decline in disability benefit inflows, which by 2005 were only 1/3 their levels in the late 1990's. But inflows into unemployment benefits and early retirement rose at the same time.

Who else participates in the assessment procedure? In recent years, private sector parties such as employers, vocational counselors, pension funds and insurance companies have been given the right and responsibility to contribute information, especially in the Netherlands and Chile. Employers have observed the individual over long periods. Vocational counselors are in the best position to understand the relationship between the medical impairment and ability to perform one's job. Pension funds and insurance companies can provide information and ask questions based on similar cases and statistical analysis. These sources offer new detailed information and counteract biased information. But usually the final decision about disability status rests with public officials. These reforms are discussed more fully in Part VIII.

How to encourage screening by self-selection. Often the individual has information about his own condition that is not readily observed by others. System procedures can encourage individuals to reveal such information. For example, in some cases applicants are required to incur some costs—such as an application fee or a waiting time without benefits before applications are permitted. This encourages self-selection—only those who consider themselves more likely to succeed will incur the costs and apply. In the U.S. a 5-month waiting period for benefits is required, during which the applicant will be disqualified if he works. This long period without income discourages applications by those who are able to work and think they have little chance of qualifying for benefits. But it also imposes hardships on those who are truly unable to work and have no other source of income.

Sick leave and anti-selection. The U.S. procedure is the converse of the situation in most European systems, where generous benefits, sometimes approaching a 100% replacement rate, are given immediately for an extended period of sick leave that precedes the disability application. In Europe the majority of disability beneficiaries start out on sick leave, where the screening procedure is much less demanding. They shift to disability once the sick leave is exhausted. After that long period of absence their work habits and skills have seriously

depreciated so they cannot readily find employment, which helps them to qualify for disability benefits. This process of easy eligibility and high replacement rates for sick leave entices into disability insurance many who would not otherwise have applied or qualified—it is the opposite of self-selection of those who are more seriously ill. This is one of the reasons why reciprocity rates are considerably higher in Europe than in the U.S.

To counter this situation and encourage positive self-selection, several European countries have been reducing the replacement rate paid during sick leave. In 1993 Sweden introduced a 1-day waiting period for sick leave. It also reduced its replacement rate to 65-80% for the first 90 days and 90% thereafter, instead of 90% for all, to impose a cost on individuals at the very beginning of this process. This reduced the number of short sick leaves (but increased the number of longer sick leaves) (Johansson and Palme 2005). Besides the tighter certification requirements already described, Poland cut the replacement rate for sick leaves after 3 months from 100% to 80%. By 2000 absenteeism fell by 1/3 and inflows into disability started to fall. In the Netherlands, replacement rates for sick leave were reduced from 100% to 70% for the second year of sick leave (OECD 2008). The Netherlands and Luxembourg have increased the employer's share of sick leave costs, to provide an incentive for them to discourage long absences by workers who are capable of working.

Intensified screening and counseling during sick leave can reduce reciprocity rates directly and also indirectly via self-selection. In Denmark, municipalities that used case management interviews had more workers transition off benefits and return to work with their previous employers (Hogelund and Holm 2005). The UK's Pathways to Work Program offers job counseling and job search assistance to those on sick leave; this seems to get women back into the labor force who otherwise might have stayed out longer (Adam, Bozio and Emmerson 2010; Bewley, Dorsett and Haile 2007). These arrangements can provide messages to the worker that the benefits from work exceed the expected benefits from disability insurance and can also provide feedback to the system about the worker's health status. However, they are expensive and require considerable time of trained case-workers to implement. They are undoubtedly more appropriate for high-income welfare states than for low and middle-income countries struggling with a scarcity of resources, especially skilled resources. But all countries that wish to cut down on the costs of their disability programs should start by tightening entry to and generosity of their sick leave benefits to cut off a common access path.

Is certification permanent? Is periodic reassessment required?

Once on the disability rolls, individuals hardly ever leave until reaching the old age pension or death. Reassessments are often possible in principle, but in practice they are rare and practically never result in disqualification (Figure 4 and Table 1). This is at odds with the fact that a person's medical condition can improve and new medical techniques may solve old medical problems—particularly for the growing group of young people who are disabled and for the psychiatric and muscular-skeletal diagnoses that now dominate disability systems. In the past, by-pass surgery has enabled people with heart disease to function normally, medication has counteracted back pain or arthritis and improved prosthetics have replaced limbs. The failure to re-evaluate these conditions is one potential source of classification error. Moreover, the absence of reassessment increases the incentive for individuals to apply for disability insurance that gives them a lifetime income, rather than for unemployment benefits that are only temporary.

A counter-acting argument is that assessments are expensive. If medical changes occur after many years of absence from the labor force, work skills have deteriorated in the meantime, offsetting the medical improvement. Moreover, periodic reassessments can, perversely, discourage work. When the time comes for a reassessment, an individual who has tried to work may find it difficult to maintain that he has lost the capacity to work. He is less likely to retain his benefit. This may deter him from trying in the first place. Perhaps for these reasons, reassessments are de facto absent in most countries and outflows (except for death or old age pension) are negligible. This means that a short run increase in inflows (say, due to a high unemployment rate in the economy) leads to a long run increase in recipients and costs.

Recent reforms have moved in the direction of treating most disabilities as a temporary condition that must be re-evaluated periodically. In 1984 Italy made almost all disability benefits partial and temporary; its contributory program went from having the highest reciprocity rate in Europe to the lowest. But half of the decline was offset by a commensurate increase in non-contributory means-tested programs (OECD 2003). During 2004-09 the Netherlands carried out a reassessment of all disabled under age 45 and 40% had their benefits cut or reduced because they were deemed able to work, at least partially. More than half of this 40% are now at work, at least part-time, while many of the others qualified for unemployment benefits or social assistance. In Poland since 1999 permanent disability grants have been eliminated. Beneficiaries must reapply every 3 years. Since then outflows from disability have increased, inflows have

decreased and reciprocity rates and costs have fallen sharply as a result of both effects. In 2001 Germany abolished the own-occupation test and made all disability grants temporary. In Australia, too, most new disabilities are now considered a temporary condition, like unemployment. Such individuals must work closely with private vocational counselors to retrain and find another job that they can do. But this does not solve the problem of a long term loss of capacity. (See Parts VII and VIII. Also see OECD 2006, 2007, 2008).

It seems appropriate to reserve the term “permanent” for only a small proportion of extreme disabilities that have little chance of improvement and to consider others as potentially temporary. Periodic reassessments of the latter group are an effective way to reduce classification errors, respond to changing medical conditions and increase outflows from the disability rolls.

Should the disabled be treated differently from other unemployed individuals?

The disabled are often “tagged” as being different from other unemployed individuals, entitled to more generous or longer-lasting benefits, because they are considered more deserving or less subject to moral hazard. However, the difficulty in correctly identifying the disabled and the growing realization that their labor supply elasticity is greater than expected is leading to a rethinking of whether such tagging makes sense. When individuals are eligible for more than one social insurance program, with differentiated benefits, they shop around and choose the benefit that is largest, easiest to get and lasts the longest. When unemployment, social assistance and early retirement plans are reformed, disability rolls may grow and conversely, when the disability program is reformed, other social insurance programs may grow. This means that reactions in other programs must be taken into account to estimate the total effect of policies on costs, employment and GDP. Therefore a caveat is in order against “tagging” the disabled for higher benefits or reforming the disability program alone. Benefit shopping and shifting are likely to make the final cost higher and cost saving lower than expected. Indeed, recent reforms are moving toward a more symmetrical treatment of the disabled, the unemployed and early retirees.

Some examples

Poland. Poland has had a dramatic decline in inflow rates, as a result of adopting some of these parametric changes. In the mid-1990’s Poland’s inflow, reciprocity and spending rates on disability were the highest of all OECD countries. 16% of the working age population received

benefits and 10 per thousand new workers flowed into the disability rolls each year. Combined with generous replacement rates, public spending on sickness and disability programs consumed almost 6% of GDP, double the OECD average (data from OECD Social Expenditure Database 2008, OECD 2003 and 2006; for summary pictures of Poland see Table 1 and Figures 1-5)

In 1999 Poland radically reformed its old age system, switching to a combination of notional and funded defined contribution plans. The disability system remained defined benefit and pay-as-you-go. However, shortly afterwards Poland greatly changed the rules for getting and staying certified as disabled. A totally disabled person must now be unable to do any job (instead of the previous own-job criterion). Stricter requirements were imposed on GPs and system doctors have the right to reassess. Replacement rates were cut. Permanent disability certification was abolished. Instead, beneficiaries must reapply every 3 years. One-third of all beneficiaries are reassessed each year and about 10% of these are disallowed. Moreover, the rules on paper are more strictly enforced, and over half of all applications are now rejected.

As a result, by the mid-2000's inflows into the main disability scheme dropped by two-thirds, sickness absences fell dramatically and outflow rates increased. Reciprocity rates have also fallen, although more slowly; the stock always adjusts slower than the flow. Spending on incapacity (sick leave plus disability) as a % of GDP was cut in half and by 2005 was at the OECD average, 2.7% (OECD 2006 and OECD Social Expenditure Database—SOCX2008--www.oecd.org/els/social/expenditure; see Figures 1-5).¹³

However, social pensions, unemployment benefits and early retirement recipients increased steeply at the same time. Tighter pre-conditions and lower benefits for early retirement in the new DC system in the future may lead to a resurgence of disability applicants. Also, work propensities have not increased among the disabled (although they may have increased among those who no longer qualify for benefits). But in the meantime, Poland illustrates the fact that parametric reforms can have dramatic effects on disability inflows and costs, if there is a political will and societal agreement to enforce them.

Sweden. Sweden further exemplifies the latitude for making parametric changes. Since the 1960's, 35-40% of each birth cohort has received disability benefits before reaching the normal retirement age (Karlstrom, Palme and Svensson 2008). The growth in sick leave in the early 2000's was particularly striking and was seen as the proximate cause of the high disability rate—many people on long term sick leave eventually switch to disability. By 2004 14% of the

working age population received some incapacity benefit and over 6% of GDP was spent on incapacity programs. To stem this tide the following parametric reforms were initiated during the past decade:

Sick leave has been limited to 1 year (previously it was unlimited) and the replacement rate while on sick leave has been reduced to 75-90% instead of 95-100%.

Employers can demand doctor's certification beginning with the first day of sick leave.

Doctors are given newly developed guidelines on how much sick leave may be warranted by different medical conditions. They must provide written justification if they authorize more time.

Disability benefits are granted only for permanent reductions in work capacity; applicants with temporary loss of capacity are eligible only for unemployment benefits or social assistance.

The disability assessment is based on ability to do any job in the labor market rather than own-job and rehabilitation services are readily available.

Special access to disability benefits for labor market and other non-medical reasons for workers over age 60 has been abolished.

Disability beneficiaries are allowed to keep wages plus benefits up to a limit (around US\$6000 per year) and to return to recipient status without a new assessment if they try working.

Employers are assisted via temporary tax credits and wage subsidies if they hire individuals who previously received sickness or disability benefits.

Administration of the sickness and disability system has become more centralized and uniform, to reduce regional disparities in procedures and outcomes.

As a result of these reforms, Sweden's absence rate from sickness and its inflow rate into disability benefits has fallen dramatically. However, the replacement rate for sickness and disability remains generous, workers who shift out of benefits and into work don't increase their income much and hence don't work much (that is, both the insurance benefit and the implicit tax on work remain high) and no party to the process has a strong direct financial interest in controlling costs. Therefore Sweden still spends 6% of its GDP on incapacity and its disability benefit recipiency rate still exceeds 10% of the working age population (OECD Social Expenditures Database 2008, OECD 2009; see Table 1 and Figures 1-5). It will probably need to do more.

We move on now to explore more fundamental structural reforms that have been made in other countries to change incentives.

VII. Reforms—2) Increasing work incentives

Recipients who have a partial capacity to work could maintain or improve their living standards if they could keep a partial benefit in addition to their partial wage. But this is inconsistent with the practice in most countries of withdrawing benefits from those who work. Work makes individuals ineligible for the insurance benefit because it indicates they still have some earning capacity. This withdrawal of benefits imposes a high “implicit tax” on work, which in turn discourages them from working even if they could.

The combination of implicit tax + explicit income tax + social security payroll tax + loss of child, family and housing allowances exceeds 70% and sometimes exceeds 100% of the wage an average disabled person might be able to earn in the UK, Denmark, Norway, Switzerland and the Netherlands (OECD 2006, 2007 and 2008).¹⁴ Tax rates start out lower for bottom-wage persons, but as their work increases at some point their allowances and social assistance are phased out so their marginal tax rate becomes very high. In the U.S. tax rates are lower initially but suddenly become 100% of the benefit, which disappears when earnings exceed the allowable amount (also see Hoynes and Moffitt 1997). This illustrates the trade-off, common in welfare programs, between the objectives of preventing poverty versus maintaining work incentives. This trade-off is especially relevant to the disabled, who are given a relatively high replacement rate while not working. The high replacement rates become a high implicit tax when benefits are withdrawn.

It is hardly surprising that few disabled individuals return to work under these circumstances; they are in a “benefit” or “inactivity trap,” unable to raise their standard of living by working more. Even if they were partially disabled to begin with, after a prolonged absence from the labor force their work habits and skills depreciate and they become fully disabled. Individuals lose the opportunity to improve their status, the economy loses the output they might have produced and the system’s finances suffer from their prolonged dependency.

Getting the disabled back to work quickly has become a key focal point of recent reforms. Many approaches have been tried but countries have been reluctant to try the most direct method—less generous benefits--which implies less insurance.

Thresholds

A common solution allows individuals to earn up to a threshold amount without loss of benefits. The implicit tax disappears while under the threshold. An increasing number of countries do this, but the threshold is generally very low. In the U.S., which has one of the highest thresholds, beneficiaries are allowed to earn up to \$1000 monthly (2011). A person who earns half the average hourly wage and works half time would stay under this threshold, but if he works full time or earns the average wage rate, he would exceed the threshold and lose his full benefit—a deterrent to work for most recipients.

Trial work periods

Once the benefit is withdrawn, the individual would ordinarily have to repeat the entire assessment procedure if he lost his job and wanted to become a disability recipient once again. Most beneficiaries would not want to take this risk. To overcome this barrier, recipients are sometimes allowed a trial work period with earnings above the threshold; they can re-enter the disability rolls without a new application if they prove unable to hold their higher-paying job. In the U.S. the right to re-enter holds for three years after benefit termination.¹⁵ However, raising the allowable threshold and introducing trial work periods have not proved very effective in enticing beneficiaries back to work (Schimmel, Stapleton and Song 2009). The UK and Australia also permit suspension of benefits for a year or two with a right to restart, and in Sweden and Denmark automatic reinstatement of benefits can occur after a much longer suspension period. But return to work for the disabled remains low. Many, undoubtedly, are simply unable to work or find work extremely uncomfortable. Others don't know about these trial work programs or don't understand or trust them. Or perhaps their benefit is almost as high as their potential wage, so there is little reason to work.

Eliminating the implicit tax

In 1981 Chile radically reformed its old age security system and its disability insurance system at the same time. In place of the previous pay-as-you-go defined benefit pension, workers are required to put 10% of their wages into a retirement savings account that is invested by a private pension fund chosen by the worker. The pension fund is required to purchase a policy for group disability and survivors' insurance that covers all its affiliates. If a worker becomes disabled, he is guaranteed a 70% replacement rate for full disability, 50% for partial disability, indexed to inflation. In either case the benefit is not lost if he works; that is, the implicit tax on work has been removed. Also, all pensioners, both disabled and able, are exempt

from the pension payroll tax. The elimination of the implicit tax and the explicit payroll tax should encourage work among disabled beneficiaries. But this system contains no mechanisms that encourages employers to accommodate disabled individuals to keep them on the job.

The net impact is difficult to test empirically because the new system also tightened access to disability status, so new-system beneficiaries are a more select group, with worse health problems (as proxied by higher mortality rates), than old-system beneficiaries (see Part VIII and James, Edwards and Iglesias 2009). This should lead them to have lower work propensities, in estimations that do not have good controls for individuals' health status. A preliminary econometric analysis of this issue, using retrospective panel data, showed that work probabilities did not fall for new-system disability pensioners despite their higher group mortality rates. Apparently the positive work incentive stemming from the elimination of the implicit tax and payroll tax was large enough to offset the negative impact on work stemming from the health composition effect (Edwards and James 2011).

Positive rewards for work--wage subsidies and supplements

A small number of countries (such as Denmark and the Netherlands) are experimenting with the idea of replacing benefits by wage subsidies or supplements for partially disabled individuals. Beneficiaries can keep the supplement in addition to the wage on a long term basis. The supplement augments the reward for work. The rationale is that work is more difficult for a disabled person, so some additional compensation is needed to induce him to do it. Moreover, his earning capacity has fallen, so work alone will not enable him to maintain his previous standard of living. The combination of wage plus subsidy enables a lower system expenditure per recipient while the individual maintains his living standard and contributes to the productive economy. At the same time, a smaller proportion of applicants can be designated as fully disabled, given a more generous benefit and not expected to work. Sometimes employers who hire the partially disabled are also eligible for subsidies to help defray any additional cost.

In the Netherlands, claimants are sorted into three groups—those with full disability (permanent loss of earning capacity of 80% or more), temporary or partial disability (capacity loss from 35 to 79%) and very little disability (<35%). Members of the first group are given a full disability benefit (over 70% of their last year's wage), with virtually no vocational rehabilitation. They are not expected to return to work. Members of the third group are expected to work (with cooperation from their employers); if they don't work they are eligible for

temporary unemployment benefits or social assistance, but not for disability pensions. However, the majority of applicants fall into the middle “partially disabled” group and much attention is given to keeping them in the labor force, with a wage supplement replacing the disability benefit if he utilizes more than half of his remaining capacity. The supplement replaces 70% of his lost capacity (e.g. if he has lost 50% of his capacity and uses the remaining 50%, he gets a supplement of $.7 \times .5 = 35\%$ and his total income is $50\% + 35\% = 85\%$ of his previous full-time wage). In contrast, if he doesn't work he gets a benefit that equals 70% of the minimum wage times his lost capacity ($.7 \times .5 = 35\%$ of the minimum wage). Thus, among individuals who previously earned more than the minimum wage, the wage supplement for work is higher than the benefit for non-work. This is the opposite of an implicit tax on work and makes it more likely that those with partial disabilities will stay in the labor force. (On the theoretical foundations for this approach, which gives less to non-workers than to workers, see Parsons 1996 and Waidman, Bound and Nichols 2003). At the same time, as described in the next section, employers were given incentives to accommodate and rehabilitate them.

However, the Dutch system also illustrates the contradiction between maintaining strong work incentives while insuring against low income—a contradiction that afflicts most social welfare programs. Partially disabled individuals who don't work get a benefit that keeps them above the poverty line. If their previous wage was low, close to the minimum, they get virtually no incremental supplement for work. And they lose their housing benefit as well as social assistance if they earn more than a minimal amount. Moreover, the incentive for marginal work beyond 50% of capacity is relatively small for everyone, since the wage supplement does not increase further while income and social security taxes rise. And an individual who works beyond 100% of capacity might lose his entire supplement. A positive effect on work for the high-earning subset of the partially disabled in the Netherlands seems likely, but at this point we have no firm evidence (OECD 2008 and 2010; also see Part VIII).

In Denmark, a fully disabled person gets a generous flat payment—70% of the country's average net earnings. Benefits for the partially disabled were replaced in 2003 by a wage subsidy scheme—flex jobs. Persons are eligible if they have a permanent impairment that makes them unable to do a normal full-time job—according to an assessment by their own GP. Instead, these individuals work part time and their partial wage is supplemented by 50-67% of the corresponding full time wage. Such a person could then work 50% time but earn 100% of his

previous wage. This is considerably higher than the wage supplement in the Netherlands (see above). An approved applicant gets a “waiting benefit” at the level of the full disability benefit until an appropriate flex job is found.

Does this scheme keep more partially disabled individuals in the labor force, receiving wages and wage subsidy instead of benefits? On the contrary, evaluation studies indicate that much of the subsidy has gone to individuals who previously were working full time for the same employer and presumably would have continued at that job if flex-jobs did not exist. Specifically, it is estimated that half of all firms involved would have hired the same workers on ordinary terms if the subsidy didn’t exist (Datta and Larsen 2007). Instead of shifting out of full benefits into partial benefits, as was hoped, many workers evidently shifted out of full work into partial work with wage subsidy. In fact, flex-jobs became so popular that the number of applicants increased faster than the supply and a waiting list grew. While on the waiting list, individuals got a full disability benefit. As a result, the number of workers on flex jobs grew rapidly, work did not increase and the inflow into benefits did not fall (OECD 2008 and 2010).¹⁶

The Danish experience illustrates a danger with wage supplements or subsidies: more people will apply for benefits and cut down on work if they can keep their partial wage in addition to their partial benefit and more will qualify if partial disability is the criterion or if the gatekeeper is lax. Thus there is a trade-off between encouraging work after a person has qualified versus discouraging application in the first place. The impact on total costs and on work depends on the relative size of the wage supplement and its targeting--whether those who get it are people who otherwise would have received no benefit with full work versus a full benefit with no work. In Denmark the wage subsidy was exceedingly generous, much more so than in the Netherlands. Moreover, the Dutch screening and monitoring mechanism, where employers play a major role, did a better job of controlling the inflow than in Denmark, where municipalities operate the system with funding from the central government. This is discussed more fully in see Part VIII.

Anti-discrimination legislation, workplace accommodations and employment quotas

All these attempts to encourage work by disabled individuals would have little effect if employers were allowed to discriminate against them in hiring or firing or refuse to make simple workplace accommodations that would compensate for the disability. Starting in Canada and the U.S., legislation passed over the last 20 years makes discrimination illegal and requires workplace adjustment for the disabled. In most European countries these workplace

accommodations are subsidized by the government. The Netherlands strongly encourages such workplace accommodations while Chile does not.

Anti-discrimination legislation makes it more difficult to fire a person whose productivity may have fallen as a consequence of his disability. Disabled persons who want to stay on the job are better able to do so. However, studies in the U.S. indicate that the inability to fire may make employers less willing to hire persons who are disabled or are in a high risk group, such as older persons with health problems. Firms may want to avoid the costs of workplace accommodations that they are required to provide or may fear the litigation costs they would incur if they tried to fire the worker in the future. Once such a person loses his job he may be unable to find another, and is likely to apply for disability benefits. Some analysts have argued that, for this reason, anti-discrimination legislation in the U.S. has reduced employment of the disabled, hurting the very group it was supposed to help. (For empirical analyses that set forth this argument see Acemoglu and Angrist 2001 and DeLeire 2000a and b. For a contrary view based on pre-existing trends and differences across states that had different legislative backgrounds, see Begle and Stock 2003, Jolls and Prescott 2005 and Houtenville et al 2009).

In many countries, firms face mandatory employment quotas of the disabled. For example, disabled persons must comprise 3-6% of all employees in Austria, France, Germany and Poland. This might increase the demand for and wage of disabled persons and therefore their incentive and opportunity to work. However, often the quotas are not filled or are filled with workers who would have been employed anyway. Usually employers can escape the quota requirement by paying a modest fee. In some cases, employers simply keep on the job their own workers who are encouraged to apply for partial disability benefits. Small firms are often exempt (OECD 2003 and 2010). If the quotas were a binding constraint they might distort the allocation of the disabled, who might end up working in places where their productivity is relatively low. This becomes an inefficient hidden tax. Giving incentives to workers and employers is a better approach.

VIII. Structural Reforms that Improve Information and Monitoring: A Larger Private Sector Role

Traditionally, disability assessments as well as benefit provision were in the hands of the public sector, financed by payroll taxes or general government revenues. The past two decades have seen movements toward a greater private role in providing information, insurance, finance and service--although the gatekeeper role, saying yes or no to applicants, continues to be carried out by public officials. The most salient private parties vary—from employers in the Netherlands to vocational counselors and job search facilitators in Australia to pension funds and insurance companies in Chile. In each case, the system is restructured so that these parties have a direct pecuniary interest in keeping costs under control and shifting beneficiaries into work rather than benefits.¹⁷

The Netherlands: employer responsibility for financing and monitoring

The Netherlands reforms began in the early 1990's, as a response to skyrocketing benefit reciprocity and costs. Disability beneficiary rates had been growing through the 1970's and 1980's, reaching 10% of the working age population by the 1990's. Costs of public disability insurance exceeded 6% of GDP (10% of wages)—almost as much as public old age spending. Through a gradual process, much of this cost has been shifted to private employers.

Presently, employers are required to finance up to 2 years of sick leave, with a 70% replacement rate, often topped up to 100% via collective agreements in the first year. During this period, besides paying for sick leave, employers must carry out workplace accommodation, contract with private providers for rehabilitation and vocational counseling of absent workers and submit periodic reports to the social insurance agency--at week 1, week 8, every 6 weeks thereafter and a big evaluation toward the end of the first and second years. If these reports don't document sufficient effort by the employer, he is required to finance a third year of sick leave; this has happened in 13% of all cases. (OECD 2008; Burkhauser, Daly and De Jong 2008). Sick leave is the major pathway to disability benefits in many European countries. The object here is to get persons with short and medium term illness back to work before they enter the long-term disability rolls. Workers are required to cooperate with these efforts in order to maintain their eligibility for sick leave and, eventually, for disability benefits.

If the individual still claims he is disabled after the sick leave period, he is assessed for remaining work capacity. Standardized medical guidelines are being developed for this purpose. As discussed in the previous section, claimants are sorted into three groups—those with a capacity loss of 80% or more, who are given a full disability benefit by the public system; those

with capacity loss of less than 35%, who don't get any disability benefit and are expected to work; and the middle "partially disabled" group, who are eligible for a wage supplement. A partially disabled individual becomes the employer's responsibility for ten years (after ten years benefits such individuals are covered by the public system). The employer must try to find a job that he can do, at the initial firm or some other firm, provide vocational rehabilitation and workplace accommodation if necessary and pay the (partial) wage plus wage supplement. Large employers self-insure while small and medium-sized employers use public or private insurance and their premiums are experience-rated.

What does this increase in employer's responsibility accomplish? Does it simply shift costs from the public to the private sector, or does it cut costs, raise output and produce better gatekeeper decisions? The potential efficiency gains come from improved information, reduced moral hazard and facilitation of work. Employers are well placed to provide detailed information on the remaining capacity of the worker since they (or their managers) observe him on a day-to-day basis. Previously employers had an incentive to bias this information and shunt to the disability rolls workers who had lost productivity or become redundant. Now they have no such incentive, since they will be bearing the cost, either directly or through experience-rated insurance. Unbiased information from the employer should enable a more accurate assessment.

In addition, employers have an incentive to get workers off sick leave as soon as possible, to monitor their cooperation and performance and to prevent them from transitioning to disability benefits. This includes finding a job that persons in ill health can do, changing the work environment if necessary and hiring vocational counselors who assist the worker in making adjustments—these costs are less than the disability benefits that the firm would otherwise have to pay. Besides their own firms, employers can help find jobs at other firms within their network; networks of employers in the same region or industry are a strong feature of the Dutch labor market. The object of these incentives is to keep partially disabled workers in the labor force, receiving part of their remuneration from wages rather than transfer payments, in order to reduce the tax costs and negative impact on labor force participation discussed earlier.

We observed earlier, in the case of Denmark, that the wage subsidy scheme enticed able workers into disability flex jobs, where they could work part time without a loss of pay. The number receiving full disability benefits did not decline but the number receiving full wages did decline. The municipalities in Denmark apparently were unable to separate out the able from the

partially disabled and employers had no responsibility or incentive to help them do so. The involvement of employers should counter this danger in the Netherlands. The employer must build a detailed record of the worker's efforts during the sick leave period and has an incentive to provide countervailing information during the assessment process that will lead questionable applications to be disallowed.

What does the evidence show? The rate of sickness absence (number of sick-person-days divided by total labor-days per year) fell from almost 6% in 2000 to 4% in 2005. Comparing rates at the turn of the century to those in 2008, inflows into disability fell from over 11 to 4 per 1000 of working age population and the stock of beneficiaries (as a percent of the working age population) declined from 9% to 8%.¹⁸ Public plus mandatory private spending on sick leave and disability benefits dropped from over 7% of GDP in 1990 to 5% in 2000 and less than 4% by 2007 (OECD 2008, 2010 and OECD Social Expenditure Database 2008).

The total beneficiary rate generally changes slower than the inflow as it is dominated by the large remaining stock. In the Netherlands, the decline in reciprocity rate was due to the falling inflow and, additionally, to the rising outflow as a reassessment of all beneficiaries under age 45 took place over 2004-09. In this process, 40% of all recipients had their benefit reduced or eliminated and more than half of these are now at work, mostly in temporary part-time jobs. As a result of this reassessment, the total number of recipients has fallen. But some of them will eventually re-enter the disability rolls or have already entered the social assistance and unemployment rolls. Our data on outcomes in 2007-08 do not capture these broader and longer-term effects.

Statistical estimates of the reasons for the sharp inflow decline focus on the extended period of employer-paid sick leave and employer responsibility for monitoring and reintegrating sick employees (OECD 2008 and 2010; also see Jehoel-Gijsbers 2007). Analysis of longitudinal administrative data indicates that some of the decline is attributable to experience-rating of insurance premiums. Employers whose premiums increased took preventative actions that reduced their disability rates in the following periods (Koning 2005). This is consistent with studies of the impact of experience rating in the unemployment benefit and workmen's compensation systems of the U.S. and Canada (see Hyatt and Thomason 1998).¹⁹ This beneficial effect, however, did not apply to small employers, who are more exposed to random events that they cannot control.²⁰

However, a complete evaluation of the reform remains to be done and many potential pitfalls are suggested. Insurance generally exhibits great economies of scale, which are foregone in the Dutch approach that breaks the system up into a separate pool for each employer. Even if competitive insurance pools and profit motives led to lower costs for employers, this may have been achieved through selection and shedding as well as greater efficiency. Employers have shifted toward the hiring of independent contractors or temporary workers for whom they do not have to provide disability insurance, and this may reduce access to job security and other private benefits that long-term workers get. Some evidence indicates that employers are reluctant to hire older workers and others with a high disability risk, just as pension funds in Chile tried to avoid these groups. Thus, although job retention is higher for those in the partially disabled category, hiring rates are slightly lower than for other workers in the same age range. Their mobility and new job prospects appear to be low once they leave their present employer so their over-all employment rate remains low (OECD 2008 and Burkhauser, Daly and de Jong 2008). This is similar to effects of the ADA in the U.S., described above. Those who remain on their job may be a select group with higher motivation and productivity or, alternatively, their current wage may exceed their productivity and should be counted, in part, as a transfer payment. In that case, total expenditures on benefits plus subsidies are higher than appears.

Public and private costs may be understated for other reasons as well. For example, how large are the spillover effects into social assistance and unemployment rolls? Have the employer's administrative and reporting responsibilities and his expenditures on rehabilitation and absenteeism management been taken into account? All these mandatory private costs are likely to be passed on to workers as a group in the form of lower wages. Disability cost variations among employers may then result in uneven wages (or job loss) for their workers. This may particularly be the case at small and medium-sized firms, some of whom may be hit by high costs because of their bad luck at hiring workers who turn out to be sick or disabled, through no fault of the employer.

To counter these unintended effects, a "no-risk policy" has been instituted under which the public system pays for the sick leave of newly-hired disabled workers (including those with less than 35% disability) for five years; a discount in the disability insurance premium applies for these workers and they do not affect the employer's experience rating. An insurance premium discount also applies to all newly hired workers over age 50 and to retained employees over age

54.5. The lower premiums for these high-risk groups means that firms are not penalized for employing them, but their insurance is being heavily cross-subsidized by low-risk groups across all employers. The public system also covers the cost of temporary workers and of young individuals (mostly diagnosed with mental illness) who have had little or no work experience. These are rapidly growing groups. Small and medium-sized firms are required to stay in the public pool, where they are subject to only limited experience rating. Altogether, more than half the new inflow into sickness and disability benefits remain in the public sector (OECD 2008). The fact that these measures were added demonstrates the complexity of such programs and suggests that problems did indeed develop. Addressing them requires on-going vigilance.

Finally, a similar system may not be transferable to low and middle income countries where the necessary technical and managerial skills are scarce, employers do not have a history of long-term employment relations with their workers or collaborative network arrangements with other firms and private insurance markets are undeveloped. Nevertheless, the basic lessons from the Netherlands—building work incentives into the disability system and involving private parties (such as employers and insurance companies) in the decision process who have a direct financial interest in providing information, controlling costs and keeping workers on the job—are transferable.²¹ In fact, a recent proposal by Autor and Duggan (2010) suggest that the U.S. adopt a variant of the Dutch system to stem the rising tide of disabled individuals who don't work. They propose that employers should be required to purchase private disability insurance (or to self-insure) to provide about two years of workplace accommodation, rehabilitation services and partial wage replacement, before public disability insurance takes over. In effect, this would add two years of private sickness-leave as a front-end to the current public U.S. system. The private insurance would be experience rated, so employers would have an incentive to quickly make workplace accommodations to keep the worker on the job and off benefits.²² The 60% replacement rate (capped at \$2500 monthly) would leave individuals with an incentive to earn wages through work. The hope is that, as a result of rehabilitation and accommodation, many of these workers would stay in the labor force and would never enter the public disability system. However, the Dutch system may be less transferable to low and middle income countries where the necessary technical and managerial skills are scarce, employers do not have a history of long-term employment relations with their workers or collaborative networks with other firms and private insurance markets are undeveloped.

Contrast with Denmark—subsidized flex-jobs run by municipalities

The flex job program in Denmark, described above, offers a sharp contrast with the Dutch scheme. In Denmark, too, applicants are sorted into the full, partial and no-disability groups, and partially disabled workers enter the flex-job program where they work part-time and get a wage supplement, as in the Netherlands. But the program has not reduced the number of regular disability beneficiaries, as was envisioned. Instead, the stock of full beneficiaries remained constant at 7% of the working age population while an additional 5% on flex jobs has now been added. So total spending on incapacity programs rose (OECD 2008 and 2010).

Why the different outcomes between the Netherlands and Denmark? The Danish scheme lacked the essential ingredient that the Dutch scheme had—one party to the transaction with good information and a strong incentive to counter the potential increased applications for subsidized positions.

In Denmark responsibility for running the program is vested in the municipalities, not the employer. The employer does not even play a formalized informational role and has no incentive to do so. The municipality was, for many years, reimbursed for two-thirds of its expenditures by the central government—a mismatch between financial responsibility and decision-making authority that weakened its incentive to hold costs down. Municipal officials have divided loyalties—to their constituents who want easy access to generous benefits and to the central government, whose object is to cut costs and encourage the partially disabled to work. The municipal officials, perhaps because they had no direct information or pecuniary interest, did not enforce the requirement that flex-jobs should go only to people who cannot hold a normal job—instead, movement out of normal jobs into part-time subsidized jobs took place. Different municipalities applied different criteria and had very different outcomes.

In addition, Denmark has no provision for periodic reassessment that would allow errors to be corrected. Once given, the benefit is permanent—in contrast to the massive reassessment that was just undertaken in the Netherlands. After several years of higher costs, Denmark has modified its system to give less reimbursement to the municipalities, greater reimbursement for employment-oriented measures and a smaller subsidy to workers, to improve incentives. But still larger structural changes may be needed.

Australia—outcome-based payments to service providers

Under the 2006 Welfare to Work reform in Australia, disability benefits go only to those who are evaluated as being permanently unable to work more than 15 hours per week. The rest are considered to have some remaining work capacity (similar to the Netherlands). Rather than entering the disability system, they are treated the same as ordinary unemployed workers, who are eligible to receive benefits for only a limited period. Thus, Australia now acknowledges the ambiguities inherent in the definition of disability. It emphasizes that those who can work should be expected to get back to work and given help toward that goal.

All new disability applicants are quickly assessed and those who qualify as unemployed are connected to a private (for-profit or non-profit) organization that provides training and personal support, to maximize the probability that they will return to work. To get unemployment benefits they are required to use these services and engage in job search activities. These service organizations are chosen through a competitive bidding process and are paid according to their success in placing disabled individuals in jobs—part of the fee is paid upfront, the rest after 13 and 26 weeks of successful job-placement. They are also audited and graded according to their outcomes, using a star-rating system. Contracts are rolled over for better performers while poor performers are eliminated and replaced by competitive tender. (The Netherlands and UK have also introduced outcome-based fees for vocational rehabilitation, but not as well developed). A similar program was used previously to place other disadvantaged unemployed groups and job placements strongly increased. It is hoped the same will happen to the disabled (Grubb 2006).

One criticism—service providers who are paid according to their outcomes may try to cream high quality clients, selecting those who will be easy to place. Or they may try to avoid localities with poor labor market conditions. To prevent this, the public agency that runs the program retains the right to randomly assign claimants to service providers and higher fees are paid on behalf of individuals with more severe disabilities, longer periods of prior unemployment or in areas with higher unemployment rates. Another criticism--the jobs may not be “high quality” jobs that retain workers and develop their skills. This is difficult to avoid, as employers may be reluctant to invest in workers whose future performance and tenure are very uncertain.

Has the program been successful? Outflow rates from disability programs have not increased, as they did in the Netherlands, since the stock of beneficiaries was not reassessed under the new rules—a political compromise needed to pass the reform. However, the tighter

eligibility standards led to a fall in disability inflow rates. This may eventually produce a decline in reciprocity rates, although until now the lower inflow has been offset by a rise in the duration of average benefit, as younger people with mental illness enter the disability rolls. Since the unemployment benefits that denied applicants get are smaller and of shorter duration than the previous disability benefits, total expenditures fell when this program was adopted.

Less attention has been paid to work incentives. It's not clear whether employment will rise or whether individuals in poor health will constantly reapply for disability or unemployment benefits. The drop in their income may induce individuals placed on unemployment instead of disability benefits to work. However, the implicit tax on wages has not been reduced, as in Chile or the Netherlands. In fact it has risen—due to a lower earnings threshold and steeper phase-out rate in the unemployment program. The implicit + explicit tax facing beneficiaries who work is 60-80%. So the marginal return to work is small, especially for those with low-wage prospects, and these individuals have an incentive to stay on benefits as long as possible. Will the supposedly short-term assistance turn into long-term support, preventing further cost-savings? The net impact remains to be seen.²³ (For more details see OECD 2007, 2010).

Chile—countervailing information from pension funds and insurance companies

As outlined above, workers in Chile are required to contribute to retirement savings accounts that are run by private pension funds and the pension funds are required to purchase group disability and survivors' insurance policies that cover all their affiliates. Until 2008 the insurance premium became part of the general expense of the pension fund. The premium depended in part on the previous claims experience of the fund. Pension funds could therefore gain an advantage if their favorable insurance experience led to premium that was lower than that of their competitors.

If a worker becomes disabled, he is guaranteed a 70% replacement rate for full disability, 50% for partial disability. Part of this benefit is financed by his own retirement account—a treatment that may discourage claims. The remainder is covered by a top-up of capital large enough so that he can purchase a lifetime annuity that equals the specified defined benefit. This top-up is financed by the group insurance policy purchased by the pension fund. But a higher claims ratio this year is likely to lead to higher premiums next year. Therefore, both pension funds and insurance companies have a strong pecuniary interest in controlling the probability of

successful disability claims. They are given an informational and advisory role in the evaluation procedure that allows them to pursue this goal.

Initial claims are evaluated by Regional Medical Boards, each made up of three doctors hired by the public regulator. The member presents his own medical tests, and the pension funds and insurance companies also have a non-voting representative at the hearing. Their Associations have organized a group of medical observers who regularly attend Board meetings, raise questions and monitor its work—a source of countervailing information to the claim. Pension funds and insurance companies can appeal approved claims to a Central Board and can participate in appeals brought by members; for this purpose they hire their own doctors to build strong appeals. Finally, they have representatives on the technical commission that determines the medical criteria for granting partial and total disability. These criteria are spelled out in a lengthy protocol manual that is used by all decision-making groups. The manual specifies the degree of disability given for each medically defined handicap. If the degree of disability exceeds 67%, the individual is considered totally disabled; if less than 50% he is not considered disabled; and if between 50% and 67% he is partially disabled.

This procedure contrasts with that in the old system, in which decisions were made by Regional Medical Commissions based on information provided by the individual's own GP, with no private participation, profit motive, medical protocols or Central Commission to assure uniform standards. Analysis of a retrospective sample of new and old system affiliates demonstrates that the disability hazard is only 20-35% as high in the new system as in the old. In 1999, by which time most workers were in the new system, age-specific disability benefit inflow rates for Chile were only 35-50% as high as in the U.S. and other OECD countries. Moreover, the cost saving has been achieved with better targeting of the severely disabled, as indicated by their subsequent mortality rates. In the old system, age-specific mortality rates were practically the same for disabled pensioners and nondisabled affiliates, while in the new system they are much higher for the disabled than for members as a whole. Apparently the new system does a better job of sorting between the disabled and non-disabled, excluding from benefits individuals with less serious medical problems—thereby reducing the number of false positives (see James, Edwards and Iglesias 2009).

Interestingly, the expertise of the pension fund representatives in dealing with the medical boards and appeals may have led to a perception of too many false negatives. Moreover,

since pension funds were not permitted to price differentiate among their affiliates according to risk category, they had an incentive to cherry-pick low-risk affiliates who were less likely to become disabled.²⁴ Also, different workers ended up paying a different implicit insurance premium depending on which pension fund they joined—which may not have seemed fair. These are some of the reasons why the system was changed in 2008 to require all pension funds to purchase a uniform disability policy in a pooled competitive auction process, rather than separate policies for each fund. Pension funds can no longer gain from having lower disability success rates than their competitors, since their insurance costs now depend on the disability experience of the entire covered labor force. As a result, they may have less incentive to invest in countervailing information in the future. Indeed, fees rose substantially in 2009, the first year of the new scheme (see James, Edwards and Iglesias 2010).²⁵

IX. Conclusions

It is useful to frame the design of optimal social insurance schemes as a choice made by prospective citizens of a society about the kind of risk they are willing to be exposed to and the insurance they are willing to pay for, before they know what their health and socio-economic status will be. They only know their own preferences about risk and income variability, for themselves and for others. The object is to think through optimal risk-avoidance arrangements from the viewpoint of the unborn, who are behind the Rawlsian “veil of ignorance” about their own individual positions (Rawls 1971).

Ex ante, all these unborn individuals face the same expected benefits and costs from a given set of rules, so they would choose the rules according to their pure insurance preferences. *Ex post*, when they have information about their health and income conditions, these rules imply different expected benefits and costs for each, hence redistribution as well as insurance. We try to isolate the system rules they would prefer on pure insurance grounds, before this is contaminated by knowledge about whether the rules will make them winners or losers. Of course, this can only be a hypothetical construct, because important information about a person’s condition in life is known from the moment of conception and even more from the moment of birth. From this theoretical perspective, different societies may make different choices because the willingness to accept risk differs across countries for cultural and historical reasons.

All countries, regardless of their risk preference, must contend with the ambiguity and non-observability of the disability state. In most insurance products that are sold on the market we can readily ascertain whether damage has occurred and if so how much. We know whether or not a person has died, whether a house has burned down and how much it will cost to rebuild, whether a car has been stolen and, if so, what is its market value. It is much more difficult to determine whether a person is functionally disabled.

In the past, and in some countries today (such as Chile), emphasis is placed simply on medical diagnosis. A person who has lost a leg gets x amount while a person who has lost the ability to hear or speak gets y . But currently most countries, especially high-income countries, place greater emphasis on functional ability. Unfortunately, except in extreme cases, there is no simple mapping from a person's measurable medical condition to his incapacity to carry out the activities of daily life, including work. This ambiguity leads inexorably to errors of inclusion and exclusion, moral hazard and the likelihood that, after implementing a scheme that seemed optimal, costs will exceed the expected amounts that citizens were willing to pay. The subjectivity of the disability state therefore diminishes the degree of insurance that is optimal, once implementation issues are taken into account. Many countries are now grappling with this dilemma—they would like to provide a high degree of insurance, as represented by a high wage replacement rate, but the costs have turned out to be non-sustainable.

The most direct approach would be to cut benefits. But this would also cut the value of the insurance, a course that countries have thus far been reluctant to follow. Instead, recent reforms have placed greater emphasis on *obtaining more accurate information and creating incentives* so that the interests of able and disabled individuals, employers and gatekeepers are aligned. Following is a summary of some of the rules that have this goal and seem to meet the ex ante Rawlsian criterion, with recent or proposed reforms italicized:

Eligibility. Eligibility for disability benefits should be based on the *incapacity to do any job*, rather than the incapacity to do one's previous job. Eligibility for benefits from the contributory scheme should be strict enough to discourage adverse selection, while those who don't meet this criterion are candidates for smaller non-contributory benefits that avoid poverty.

Assessment and reassessment. The assessment of disability status should be carried out by a *team of doctors and vocational experts hired by the insurance system*. While the individual's own GP should provide information, an independent unbiased evaluation by experts

who have a broader frame of reference and knowledge of health-work relationships is essential. The assessment should be based on *written guidelines* that are uniformly applied. Appeals of the initial decision should be two-sided—*both denied and approved claims should be appealable*, in order to arrive at the most accurate ruling. Certification should be granted only for a limited period of time. Beneficiaries should be carefully *reassessed on a regular basis*, to take account of changes in the individuals' medical condition and medical advances that make a previous condition less disabling.

Self-selection. Expect disabled workers to bear some cost—a waiting period, lower replacement rate, a requirement that he show up for retraining or job search activities, to discourage healthier workers from applying. In systems with retirement savings accounts, the balance in the account could be applied toward disability benefits with the same object.

Permanent and full vs. partial and temporary disability. Criteria should be set such that only a relatively small group is considered fully and permanently disabled. *Most approved applicants should be considered partially or temporarily disabled*, able to carry out some work and rewarded for doing so. In some cases, retraining will be needed and should be provided.

Work incentives. *Benefits for the partially disabled should be modest and augmented by income from work.* To encourage work and raise their standard of living, partially disabled beneficiaries should be *permitted to keep their wage in addition to their disability benefits*. The implicit tax on work might even be replaced by a work subsidy, paying a low benefit to those who don't work and a higher benefit to those who do—but this requires careful targeting.

Family, child and family allowances. *Special care should be taken in granting family, child and housing allowances that are phased out* when the recipient works. These can produce an implicit+explicit tax on work by beneficiaries that exceeds 100%. One possibility is to give a relative modest grant, based on calculations of the expected earnings capacity of the disabled person and his family. The allowances can then be left in place when these earnings are realized.

Sick leave. Benefit generosity and screening for sick leave should be set at levels that *discourage extended sick leaves*, which often become an easy entry to long term disability benefits. *A doctor's certificate* should be required after a short period of sick leave.

Local vs. central administration. If the system is administered locally, a substantial portion of costs should also be borne locally, to *align their interests with that of the system*. Care should be taken to *avoid large regional disparities*.

Private sector role. An important role should be built in for *participation by private sector parties who have a direct interest in controlling costs and keeping individuals on the job.* Particularly relevant are employers, who are well positioned to observe the individual and make workplace accommodations and insurance companies, which are experienced in gathering and interpreting aggregate data and bringing them to bear on individual cases. They can play important informational and financing roles.

Coordination with other social insurance programs. Benefit levels and eligibility for disability benefits, unemployment compensation and early retirement *pensions should be coordinated*, so attempts to reform one program don't result in large spillovers into the other.

The most promising reforms focus on 1) giving disabled individuals incentives to work and 2) involving private sector parties, who have a pecuniary interest in the outcome, in the assessment of disability and the financing of disability insurance. In contrast to traditional systems that withdraw benefits once the individual earns beyond a low threshold, Chile allows beneficiaries to keep their wage in addition to their benefit. Denmark and the Netherlands go further in this direction. Their benefit to partially disabled persons takes the form of a wage supplement or subsidy, which the individual gets only if he works. A lower benefit can be paid, because it will be augmented by a partial wage. Both the disabled individual and the broader economy gain if these incentives encourage beneficiaries to work.

One danger of the Dutch and Danish approach is that more workers may be enticed into applying for disability benefits, to get the wage subsidy that turns part-time work into full-time pay. To avoid these false positives and keep costs controlled more generally, a reliable source of information is needed that deters claims by those who have the capacity to work normally. The Chilean scheme gives pension funds and insurance companies the power and incentive to participate in the assessment process, present information, contest claims and bring appeals, thereby constraining false positives. The Dutch scheme gives employers the responsibility to help evaluate the capacity of individuals who are on sick leave or applying for disability benefits, to find a job that they can do and to monitor their cooperation and performance. Since employers pay for sick leave and partial disability benefits, either directly or through experience-rated insurance premiums, they have an incentive to contain successful claims and make reasonable workplace accommodations to keep individuals as workers rather than beneficiaries. The Danish system, which did not include this mechanism for controlling the number of beneficiaries, was

inundated with applicants for subsidized flex-jobs. Australia attempts to wean people off dependency without wage subsidies, keeping a high implicit tax, by making them eligible only for temporary unemployment benefits and giving them extensive vocational counseling. Counselors, who are paid on the basis of outcomes have an incentive to conduct successful training and job search programs—another use of private sector initiatives. But beneficiaries with low earning prospects have an incentive to stay on benefits as long as possible. These reforms are promising but many questions remain and a careful evaluation has yet to be done.

Finally, the arrangements that work well in high-income countries may not work well in low and middle-income countries where much employment is small scale and transient and where workers move in and out of the informal sectors. To run a disability system efficiently requires considerable administrative costs and a large pool of personnel who have the skills to manage complex programs and make difficult judgments—but this resource is scarce in most low and middle-income countries. Ambiguous eligibility criteria lend themselves to mismanagement and corruption. Chronic unemployment and low educational and income levels in an economy make it problematic to give the medically disabled special treatment. For these reasons, low and middle-income countries should proceed cautiously with their disability programs—keeping them small, benefits modest and eligibility tight, closely tied to objective medical diagnoses, while laying the groundwork for structures and procedures that will avoid poor information, perverse incentives, high costs and low labor force participation in the future.

Table 1: Characteristics of Traditional Disability Schemes and Recent Reforms

Country (X if large private sector role)	Inability to do own job or any job?	Assessed by own or system doctor?	Benefits for partial disability?	Permanent/temporary? Re-evaluation?	Replacement rate, av. earner, full disab.	Benefits lost or gained if work after threshold
Australia X	Any job	GP + work capacity assessment	Treated as unemployed if can work >15 hrs wkly	Permanent and temporary. Reassessment possible but rare	Flat--25% of average earnings	Benefits lost
Denmark	Any job Capacity loss, resid. capacity	GP + specialist	flex-jobs	Permanent, no retests	70-90%-sickness; 66%+child ben--full dis	On-going flex job subsidy; but child, housing & social assistance lost
Netherlands X	Any job Capacity loss and residual capacity	GP + system + company doctors + rehab+ med guidelines	Yes, more put into partial group	Permanent and temporary; reassess all <45	70-100%-sickness; 72%+child ben.-- full disability	On-going wage supplement; but child, housing & social assistance lost
Norway	Any (slow implementation)	GP is primary	Yes but most are full dis.	Both but mostly permanent; few re-evaluations	100%--sickness 70%+child ben.--dis	Benefits & child supplement lost
Poland	Total dis-any job Partial dis--usual	GP+stricter reporting requiremts+ system doctors	yes	All temporary; must reapply after 3 years	80-100%--sickness; 50%+child ben.--dis.	Benefits & child supplement lost; ltd. wage subsidy to employers
Sweden	Any job	GP+medical guidelines	yes	Permanent but subject to review	70-90%	Benefits lost; tax credits & wage subsidies for dis. workers & empl.
Switzerland X	Any job Earning capacity	GP + insurance doctors+ regional med service	Yes but most are full disability	Temporary or permanent; ad hoc re-evaluations	80-100%--sickness 70%+child ben.--dis.	Benefits & child supplements lost
UK	Sick leave-own job Dis ben-Any job	GP + personal capacity assessment	no	Benefits set for fixed period, then reassessed	Flat--25% of av. wage + child ben.	Benefits + child support lost; but tax credit for low earning dis.
U.S.	Any job	GP is primary	no	De facto permanent	60%	Benefits lost; trial work period
Chile X	Medical criteria	GP, system, pension fund and ins. co. doctors + med manual	Yes but most are full dis.	Permanent	70%	Benefit is independent of work

Source: OECD 2003, 2006, 2007, 2008, 2009, 2010. Replacement rate for U.S. is from SSA, "Trends in the Social Security and Supplementary Security Income Disability Programs", for median worker, based on unpublished data, mid 2000's (http://www.ssa.gov/policy/docs/chartbooks/disability_trends/sect05.html). Chile data are from James, Edwards and Iglesias 2009. Reforms are in bold. X means private employers, pension funds and/or providers play large mandatory informational, financing or vocational service roles. Capacity or earnings loss takes into account education and experience, albeit in a loose way.

Table 2: Key Policy Design Issues

- 1. Definition.** What is the definition of disability?
- 2. Eligibility.**
 - What are the eligibility conditions for the contributory scheme?
 - How many months of contributions required?
 - Is the criterion inability to do one's own job, usual job or any job?
 - Any special eligibility rules for older workers?
 - Is mental disability covered?
- 3. Partial vs. full disability.**
 - What % incapacity is needed for full benefits?
 - Are the partially disabled eligible for benefits?
 - What % incapacity is needed for partial benefits?
 - Is there a written protocol for determining degree of incapacity?
 - Are the partially disabled expected to work? Do they lose benefits if they work?
 - Are retraining and rehabilitation required? encouraged? subsidized?
- 4. Non-contributory disability programs.**
 - Is there a non-contributory disability program?
 - Is it residence-based? flat benefit? means-tested?
- 5. Assessment procedure.**
 - Is the medical assessment done by the person's own GP or by the system's doctor?
 - By a team or an individual?
 - Are GPs given medical guidelines on relationship between health and ability to work?
 - Do any of the following participate or provide information in the assessment and monitoring procedure-- employers, vocational counselors, pension funds or insurance companies?
 - Do applicants have a waiting period without work or benefits?
 - Who is the benefit gatekeeper or ultimate decision-maker?
- 6. Appeals.**
 - What is the appeals procedure?
 - Which parties can appeal initial decisions?
 - Can approved claims (as well as denials) be appealed? By whom?
- 7. Paid sick leave.**
 - Is there a period of paid sick leave prior to disability claims? If so, how long? Who pays?
 - Is a doctor's certificate needed? Are doctors given guidelines on usual sick leave needed for various illnesses?
 - Are workers given counseling or other support on returning to work after sick leave?
- 8. Reassessment.**
 - Is disability certification temporary or permanent?
 - Is periodic reassessment required on paper? Is it carried out in practice? How often?
 - Is it the same procedure as the initial assessment?
- 9. Replacement rate and other benefits.**
 - What is the usual replacement rate for an average fully disabled individual?
 - Is the replacement rate higher for low earners?
 - Is the disability benefit larger or smaller than the average unemployment benefit?

Is it larger or smaller than the average early retirement pension?
What is the usual replacement rate for an average partially disabled individual?
Is the partial replacement rate higher for low earners?
What is the replacement rate during sick leave? Does it start from day 1?
How large is the non-contributory disability benefit?
Do disability beneficiaries in the contributory or non-contributory programs get any of the following: a housing allowance? a family allowance? a child allowance? any special medical benefits? any allowance for caregivers and special equipment needed?

10. Work (dis)incentives.

Are benefits reduced or eliminated if the beneficiary earns wages?
At what rate are benefits phased out as earnings rise?
At what level of earnings does the phase-out end because benefits become 0?
Are housing, family or child allowances also phased out?
Is there a threshold of earnings below which benefits can be retained? If so, what threshold?
Is there a trial work period during which a person can return to benefits if work fails? If so, how long is the trial work period?
Is there a wage supplement or subsidy for partially disabled individuals who work? How large is it? Does it increase with amount worked?

11. Employer costs and responsibilities.

Is the employer required to finance sick leave? How long?
Does he manage and report on sick leave absences?
Is he required to insure and/or finance disability benefits? If so, for how long?
Is he required to provide information to the gatekeeper about the individual's capacity to work? To monitor the worker after return to work?
Does he have to find a job for disabled workers?
Are employers required to make workplace accommodations for their disabled employees?
Are these accommodations subsidized?
Are employers prohibited from discriminating against the disabled in hiring and firing?
Do they face quotas for employing disabled workers?
Do employer costs rise if their workers become disability recipients?
Do their net costs rise or fall if their disabled employees work?

12. Pension funds and insurance companies.

Do private pension funds and/or insurance companies provide mandatory or quasi-mandatory sick leave or disability benefits?
Do they have an incentive to keep individuals off the disability rolls?
Do they provide information or have any other role in the assessment process?

13. Vocational counselors.

Are private vocational counselors involved in the mandatory system?
Are they paid according to # of clients or success in placing them?

14. Central vs. local administration.

Is the public system administered by the central government or by local or provincial governments?
Are costs paid by the central government, local governments or social security system?
What procedures exist for coordination?

Table 3: Public and Mandatory Private Spending on Incapacity as % of GDP (compared with old age spending), OECD Countries, 2005

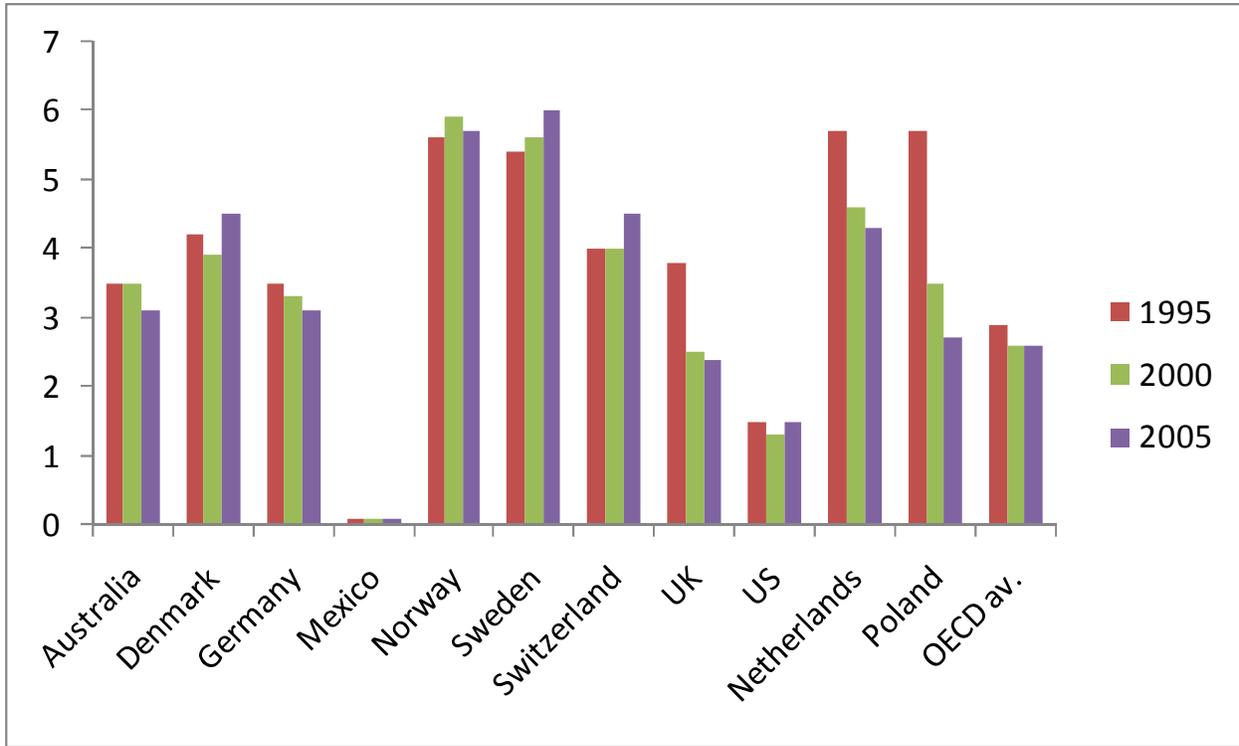
country	Incapacity			Old age			Incap/OA total
	public	Mand. pvt	total	public	Mand pvt	total	

This table will be filled in later, containing many countries, from OECD data base.

Table 4: Outcomes to be monitored (last year and changes over last decade)

1. What are the total public expenditures on disability benefits (as a % of GDP)?
2. What are the mandatory private expenditures (if any) on disability benefits?
3. What is the new inflow into disability rolls?
4. What is the total outflow? To death? To old age pension? To work? Other?
5. What is the total recipiency rate, as a % of the working age population?
6. What proportion of disability benefit recipients work?
7. What proportion of beneficiaries receive partial benefits?
8. What % of claims are denied?
9. What % of denied claims and approved claims are appealed?
10. In what % of appealed claims were the initial decisions over-ruled?
11. What proportion of recipients had their cases reassessed each year?
12. What proportion of beneficiaries have their benefits cut or eliminated during reassessment?
13. Have any of these numbers changed over the past decade, as a result of changing economic or demographic circumstances? Policy reforms?

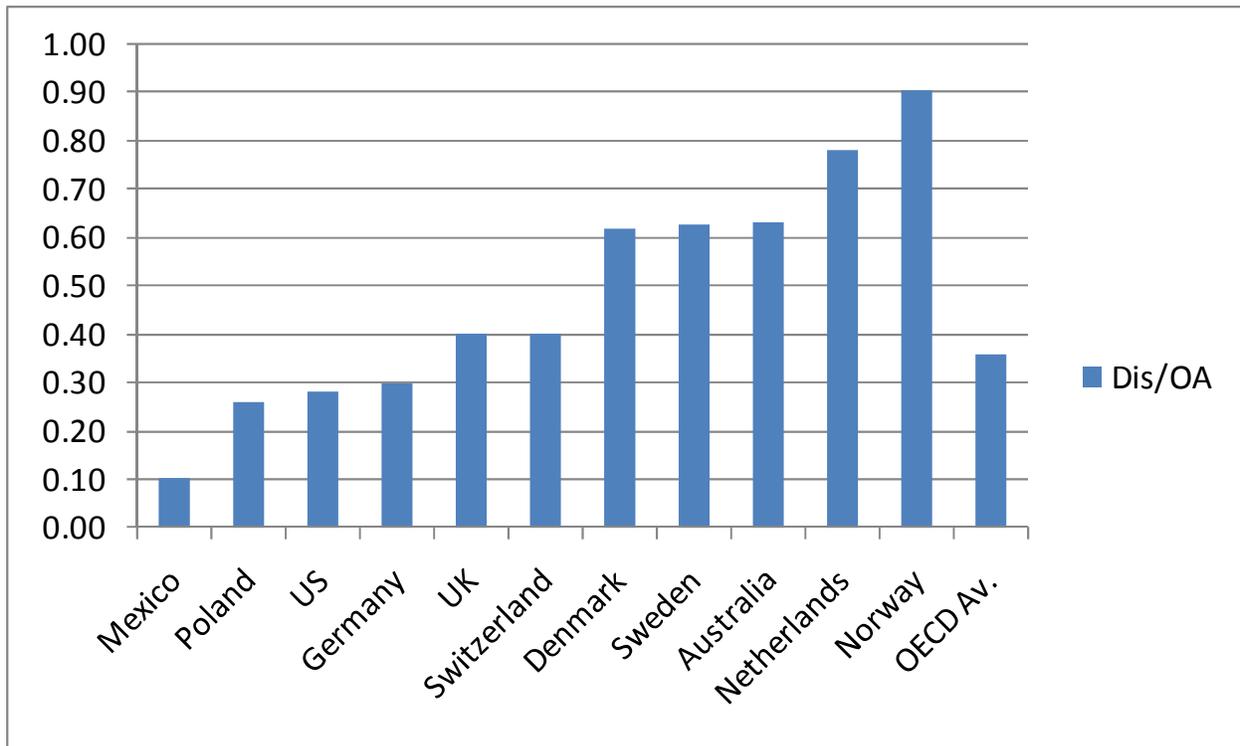
Figure 1: Spending on Incapacity as % of GDP, 1995, 2000, 2005



Source: OECD Social Expenditure Database 2008 (SOCX, www.oecd.org/els/social/expenditure)

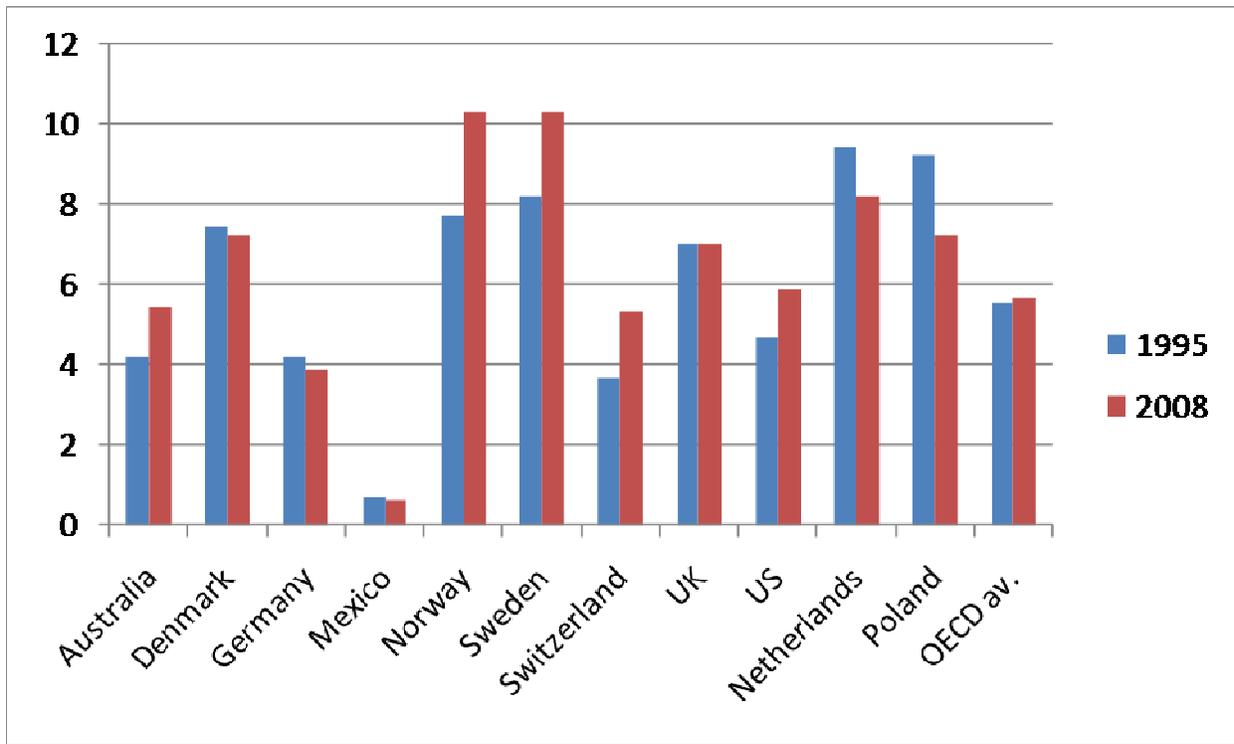
Note: countries with large recent drops are on right.

Figure 2: Public + Mandatory Private Spending on Disability/Old Age Spending, 2005



Source: OECD Social Expenditure Database 2008 (SOCX, www.oecd.org/els/social/expenditure)

Figure 3: Disability benefit recipients as % of population age 20-64, 1995 and 2008

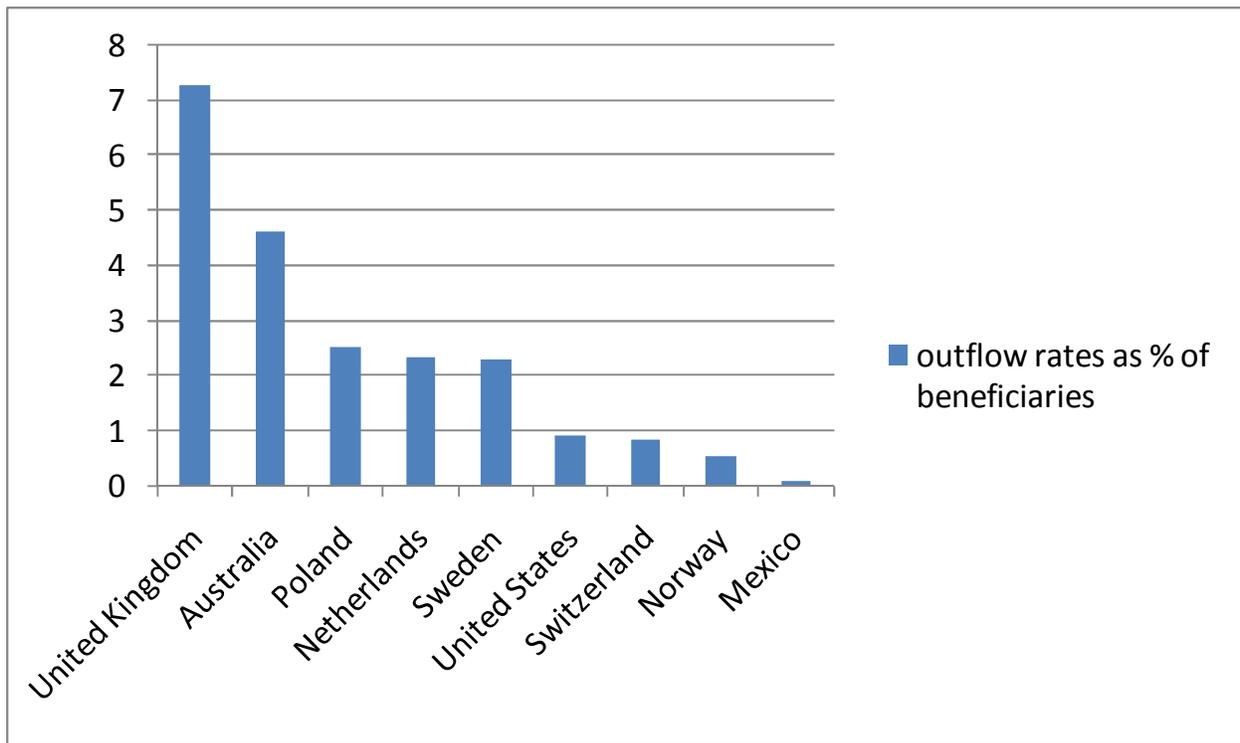


Source: OECD 2010. www.oecd.org/els/disability

Note: Latest year for Poland and UK is 2007

Countries with large recent drops are on right.

Figure 4: Outflow Rates as % of All Disability Beneficiaries, mid 2000's

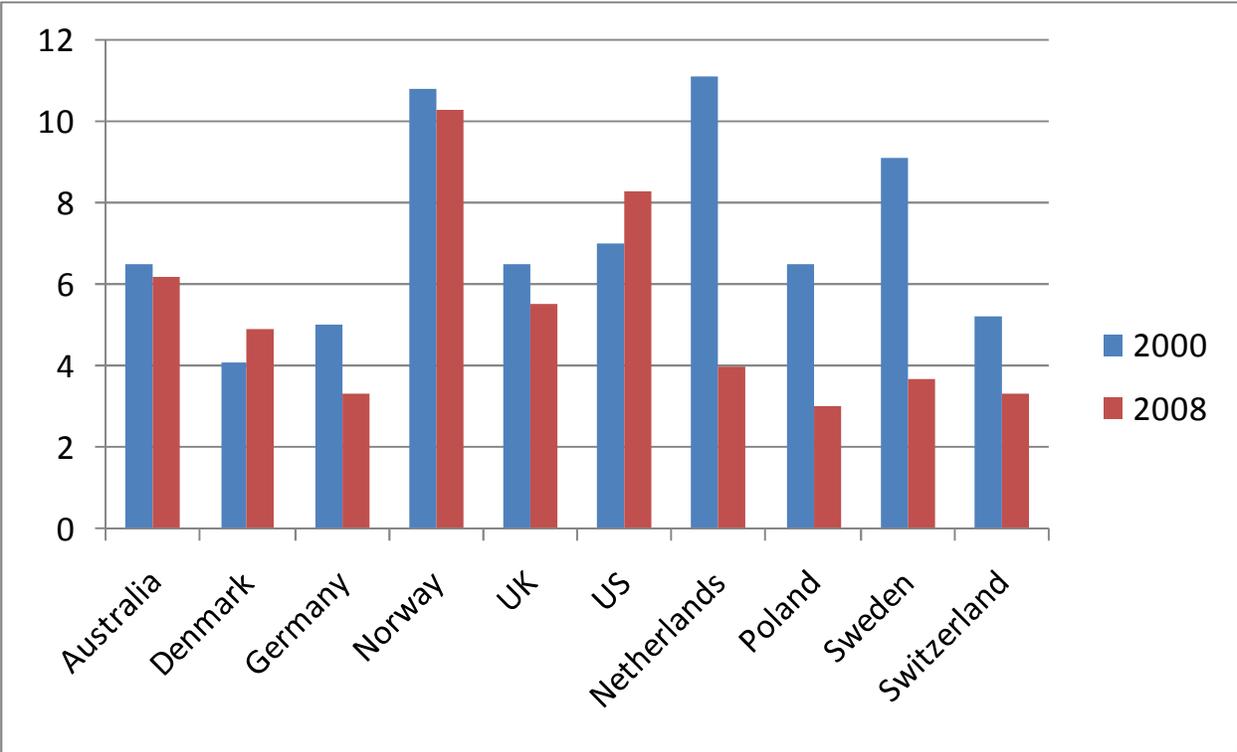


Source: OECD 2010. www.oecd.org/els/disability

Note: These numbers give the proportion of benefit recipients who leave the rolls each year for employment voluntarily or because they have been found ineligible or for other reasons besides death or receipt of old age pension.

Disability period starts after varying periods of sick leave. In the UK, mandatory sick leave period is shorter than in other European countries so disability recipients include many with short term incapacities who are more likely to return to employment. This is the major reason for the relatively high outflow rate in the UK.

Figure 5: Inflow to disability benefits per 1000 of working age population, 2000 and 2008



Source: OECD 2010 and 2007. www.oecd.org/els/disability

References

- Aarts, Leo and Philip R. de Jong. 1992. *Economic Aspects of Disability Behavior*. Amsterdam: North-Holland.
- Acemoglu, Daron and Joshua Angrist. 2001. "Consequences of Employment Protection? The Case of the Americans with Disabilities Act." *J.P.E.* 109(5):
- Adam, Stuart, Antoine Bozio and Carl Emmerson. 2010. "Reforming disability Insurance in the UK: Evaluation of the Pathways to Work Programme." London: Institute for Fiscal Studies.
- Angrist, Joshua, Stacey Chen and Brigham Frandsen. 2009. "Did Vietnam Veterans Get Sicker in the 1990's? The Complicated Effects of Military Service on Self-Reported Health." NBER Working Paper #14781.
- Autor, David and Mark Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *Q.J.E.* 118:157-205.
- Artur, David and Mark Duggan. 2006. "The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding." *Journal of Economic Perspectives.* 20(3): 71-96.
- Artur, David and Mark Duggan. 2010. "Supporting Work: A Proposal for Modernizing the U.S. Disability Insurance System." Discussion paper. Washington DC: The Center for American Progress and The Hamilton Project.
- Banks, James, Arie Kapteyn, James Smith and Arthur van Soest. 2008. "Work Disability is a Pain in the *****". in *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly*. ed. Cutler, David and David Wise. Chicago: University of Chicago Press, 251-294.
- Begle, K. and A. Stock. 2003. "The Labor Market Effects of Disability Discrimination Laws." *Journal of Human Resources.* 38:806-59.
- Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, John Rust and Sofia Sheidvasser. 1999. "An Empirical Analysis of the Social Security Disability Application, Appeal and Award Process." *Labour Economics.* 6: 147-178.
- Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, Sofia Cheidvasser and John Rust. 2004a. "How Large is the Bias in Self-Reported Disability?" *Journal of Applied Econometrics.* 19(6): 649-70. (also NBER WP#7526).
- Benitez-Silva, Moshe Buchinsky and John Rust. 2004b. "How Large are the Classification Errors in the Social Security Disability Award Process?" Cambridge MA: NBER WP# 10219.
- Bewley, Helen, Richard Dorsett and Getinet Haile. 2007. "The Impact of Pathways to Work." Research Report No. 435, London: Dept for Work and Pensions.

- Black, Dan, Kermit Daniel and Seth Sanders. 2002. "The Impact of Economic Conditions on Participation in Disability Programs: Evidence from the Coal Boom and Bust." *American Economic Review*. 92(1): 27-50.
- Boadway, Robin, Nicolas Marceau and Motohiro Sato. 1997. "An Agency Model of Welfare and Disability Assistance." Queen's Economics Department Working Paper No. 949. Queen's University, Ontario, Canada.
- Borsch-Supan, Axel. 2010. "Work Disability: The Effects of Demography, Health and Disability Insurance." in *Research Findings in the Economics of Aging*, ed. David Wise. Chicago: Univ of Chicago Press, pp. 37-62.
- Bound, John. 1989. "The Health and Earnings of Rejected Disability Applicants." *American Economic Review*. 79, 482-503.
- Bound, John. 1991. "The Health and Earnings of Rejected Disability Insurance Applicants: Reply." *American Economic Review*. 81(5): 1427-34.
- Bound, John and Richard Burkhauser. 1999. "Economic Analysis of Transfer Programs Targeted on People with Disabilities." *Handbook of Labor Economics*. 3: 3417-3528.
- Bound, John, Julie Berry Cullen, Austin Nichols, and Lucie Schmidt. 2004. "The Welfare Implications of Increasing Disability Insurance Benefit Generosity." *Journal of Public Economics*. 88(12): 2407-2514.
- Burkhauser, Richard, J.S. Butler, and Robert Weathers II. 2001-02. "How Policy Variables Influence the Timing of Applications for Social Security Disability Insurance." *Social Security Bulletin*. 64(1): 52-83.
- Burkhauser, Richard, Mary Daly and Philip de Jong. 2008. "Curing the Dutch Disease: Lessons for United States Disability Policy." Ann Arbor MI: MRRC working paper WP2008-188.
- Cai, Lixin, Ha Vu and Roger Wilkins. 2007. "Disability Support Pension Recipients: Who Gets Off (and Stays Off) Payments?" *The Australian Economic Review*, 40(1): 37-61.
- Campolieti, Michele. 2002. "Moral Hazard and Disability Insurance: On the Incidence of Hard-to-Diagnose Medical Conditions in the Canada/Quebec Pension Plan Disability Program." *Canadian Public Policy XXVIII* (3). Pp. 419-441.
- Campolieti, Michele. 2003. "Disability Insurance Eligibility Criteria and the Labor Supply of Older Men." *Economics Bulletin*. 10:3, pp. 1-7.
- Campolieti, Michele. 2004. "Disability Insurance Benefits and Labor Supply: Some Additional Evidence." *Journal of Labor Economics*. 22: 4, 863-889.

- Chandra, Amitabh and Andrew Samwick. 2008. "Disability Risk and the Value of Disability Insurance." in *Health at Older Ages*. ed. David Cutler and David Wise. Chicago: Univ Chicago Press, pp. 295-336.
- Chen, Susan and Wilbert van der Klaauw. 2008. "The Work Disincentive Effects of the Disability Insurance Program in the 1990's." *Journal of Econometrics*. 142(2): 757-784.
- Cullinan, B. Gannon and S. Lyons. 2010. "Estimating the Extra Cost of Living for People with Disabilities." *Health Economics*. (online DOS: 10.1002/hec.1619).
- Cutler, David, Mary Beth Landrum, and Kate Stewart. 2009. "How do the Better Educated Do it? Socioeconomic Status and the Ability to Cope with Underlying Impairment." in *Developments in the Economics of Aging*. ed. Wise, David. Chicago: University of Chicago Press. 203-254.
- Datta, Gupta, N. and M. Larsen. 2008. "Evaluating Employment Effects of Wage Subsidies for Disabled Individuals—The Danish Flexjobs Scheme." Copenhagen: Danish National Institute of Social Research.
- deJong, Philip, Maarten Lindeboom and Bas van der Klaauw. 2006. "Screening Disability Insurance Applications." Centre for Economic Policy Research (CEPR) DP5564 and IZA DP 1981, Bonn.
- DeLeire, Thomas. 2000a. "The Unintended Consequences of the Americans with Disabilities Act." *Regulation* 23(1).
- DeLeire, Thomas. 2000b. "The Wage and Employment Effects of the Americans with Disabilities Act." *Journal of Human Resources*. 35(4): 693-715.
- Diamond, Peter and Eytan Sheshinski. 1995. "Economic Aspects of Optimal Disability Benefits." *Journal of Public Economics*. 57, 1-23.
- Duggan, Mark, Perry Singleton and Jae Song. 2005. "Aching to Retire? The Rise in the Full Retirement Age and its Impact on the Disability Rolls." NBER Working paper #11811
- Duggan, Mark and Scott Imberman. 2008. "Why are the Disability Rolls Skyrocketing? The Contribution of Population Characteristics, Economic Conditions and Program Generosity." in *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly*. ed. Cutler, David and David Wise. Chicago: University of Chicago Press. 337-80.
- Edwards, Alejandra Cox and Estelle James. 2011. "Impact of Chile's Pension System on Work Propensities of Men and Women: Evidence from a Retrospective Data Set." Draft ms.
- French, Eric and Jae Song. 2009. "The Effect of Disability Insurance Receipt on Labor Supply." Chicago: Federal Reserve Bank of Chicago Working Paper Series. WP 2009-05.

- Goldblatt, Beth. 2009. "Gender, Rights and the Disability Grant in South Africa." *Development South Africa*, 26(3) 369-82.
- Golosov, Mikhail and Aleh Tsyvinski. 2006. "Designing Optimal Disability Insurance: A Case for Asset Testing." *J.P.E.* 114, 257-79.
- Gruber, Jonathan and Jeffrey Kubik. 1997. "Disability Insurance Rejection Rates and the Labor Supply of Older Workers." *Journal of Public Economics*: 1-23.
- Gruber, Jonathan. 2000. "Disability Insurance Benefits and Labor Supply." *Journal of Political Economy*. 1162-1183.
- Grubb, David. 2006. "Australia's Quasi-market Delivery of Case Management." Presented at ISSA workshop on Unemployment Insurance and Employment Maintenance. Brussels.
- Haveman, Robert and B. Wolfe. 1984. "Disability Transfers and Early Retirement: A Causal Relationship." *Journal of Public Economics*, XIV, 47-66
- Haveman, Robert and Barbara Wolfe. 2000. "The Economics of Disability and Disability Policy." in *Handbook of Health Economics*. Vol. 1, ed. A.J. Culyer and J.P. Newhouse. Elsevier Science B.V.
- Hasselius, Patrik and M. Persson. 2007. "Incentive and Spill-over Effects of Supplementary Sickness Compensation." IFAU Working Paper. No. 2007:16. Institute of Labour Market Policy Evaluation. Uppsala, Sweden.
- Hasselius, Patrik, Per Johansson and Johan Vikstrom. 2008. "Monitoring and Norms in Sickness Insurance: Empirical Evidence from a Natural Experiment," IFAU Working Paper No. 2008:8, Institute for Labour Market Policy Evaluation, Uppsala, Sweden.
- Hasselius, Patrik, Per Johansson and Laura Larson. 2005. "Monitoring Sickness Insurance Claimants: Evidence from a Social Experiment." Working Paper Series No. 2005:15, Institute for Labour Market Policy Evaluation (IFAU), Uppsala, Sweden.
- Høgelund, Jan and Anders Holm. 2005. "Case Management Interviews and the Return to Work of Disabled Employees." *Journal of Health Economics*. 25: 500-519.
- Houtenville, Andrew, David Stapleton, Robert Weathers II and Richard Burkhauser, ed. 2009. *Counting Working-Age People with Disabilities: What Current Data Tell us and Options for Improvement*. Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research.
- Hoynes, Hillary and Robert Moffitt. 1997. "Tax Rates and Work Incentives in the Social Security Disability Insurance Program: Current Law and Alternative Reforms." NBER WP 6058.
- Huttunen, Kristiina, Jukka Pirttila and Roope Uusitalo. 2010. "Employment Effects of Low-Wage Subsidies." Munich: CESifo Working Paper No. 3043.

Hyatt, Douglas and Terry Thomason. 1998. "Evidence on the Efficacy of Experience Rating in British Columbia." *A Report to the Royal Commission on Workers' Compensation in British Columbia*.

James, Estelle and Augusto Iglesias. 2008. "Disability Insurance with Pre-Funding and Private Participation: the Chilean Model." in *Funded Systems: their Role in Solving the Pension Problem*. FIAP. Santiago, Chile. (longer versions available as World Bank Social Protection Discussion Paper 0801 and MRRC Working Paper 2006-111).

James, Estelle, Alejandra Cox Edwards and Augusto Iglesias. 2009. "The Impact of Private Participation and Countervailing Information on Disability Costs: Evidence from Chile." *Journal of Pension Economics and Finance*. 8(3): 291-320.

James, Estelle, Alejandra Cox Edwards and Augusto Iglesias. 2010. "Chile's New Reforms." Report No. 32. National Center for Policy Analysis. Dallas, Texas.

Jehoel-Gijsbers, G. 2007. "Beter aan het werk. Trendrapportage ziekteverzuim, arbeidsongeschiktheid en werkhervatting." The Hague: Social and Cultural Planning Bureau.

Johansson, Per and M. Palme. 2005. "Moral Hazard and Sickness Insurance: Empirical Evidence from a Sickness Insurance Reform in Sweden." *Journal of Public Economics*. 89(9-10): 1879-90.

Johansson, Per and M. Palme. 2002. "Assessing the Effect of Public Policy on Worker Absenteeism." *Journal of Human Resources*. 37(2): 381-409.

Jolls, Christine and J.J. Prescott. 2005. "Disaggregating Employment Protection: The Case of Disability Discrimination." Harvard Public Law Working Paper #106. Cambridge, MA: Harvard Law School.

Jones, Andrew and own O'Donnell. 1995. "Equivalence Scales and the Costs of Disability." *Journal of Public Economics*. 56: 273-289.

Kangasharju, Aki. 2005. "Do Wage Subsidies Increase Employment in Subsidized Firms?" *Economica*. 74: 51-67.

Kaptyen, Arie, James Smith, Arthur van Soest and James Banks. 2010. "Labor Market Status and Transitions during the Pre-retirement Years: Learning from International Differences". in *Research Findings in the Economics of Aging*. David Wise ed. Chicago: Univ Chicago Press.

Karlstrom, Anders, Marten Palme and Ingemar Svensson. 2008. "The Employment Effect of Stricter Rules for Eligibility for DI: Evidence from a Natural Experiment in Sweden." *Journal of Public Economics*. 92: 2071-82.

Koning, Pierre. 2005. "Estimating the Impact of Experience Rating on the Inflow into Disability Insurance in the Netherlands." CPB Discussion Paper No. 37, The Hague.

Lee, J.-K. 2005. "Evaluation of and Lessons from Wage Subsidy Programmes in OECD Countries." In *Employment Insurance and Public Employment Services in Selected Countries*. ed. Jaeho Keum. Korea Labor Institute.

Leonard, Jonathan. 1986. "Labor Supply Incentives and the Disincentives for Disable Persons." in *Disability and the Labor Market: Economic Problems, Policies and Programs*. Ed. M. Berkowitz and M.A. Hill, Ithaca, NY: ILR Press.

Leonard, Jonathan. 1979. "The Social Security Disability Program and Labor Force Participation." NBER WP #392.

Low, Hamish and Luigi Pistaferri. 2010. "Disability Risk, Disability Insurance and Life Cycle Behavior". NBER WP #15962.

Maestas, Nicole and Na Yin. 2008. "The Labor Supply Effects of disability Insurance work Disincentives: Evidence from the Automatic Conversion to Retirement Benefits at Full Retirement Age." Ann Arbor: Michigan Retirement Research Center WP #2008-194.

Maestas, Nicole, Kathleen Mullen and Alexander Strand. 2010. "Does Disability Insurance Receipt discourage work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." Ann Arbor: Michigan Retirement Research Center WP #2010-241.

Martin, John P. and David Grubb. 2001. "What Works and for Whom: A Review of OECD Countries' Experiences with Active Labour Market Policies". Working Paper Series No. 2001:14. Institute for Labour Market Policy Evaluation (IFAU). Uppsala, Sweden.

Nagi, Saad. 1969. *Disability and Rehabilitation: Legal, Clinical and Self-Concepts of Measurement*. Columbus, OH: Ohio State University Press.

OECD. 2003. *Transforming Disability into Ability: Policies to Promote Work and Income Security for Disabled People*. Paris: OECD.

OECD. 2006. *Sickness, Disability and Work: Breaking the Barriers*. Vol. 1 (Norway, Poland and Switzerland). Paris: OECD.

OECD. 2007. *Sickness, Disability and Work: Breaking the Barriers*. Vol. 2 (Australia, Luxembourg, Spain and the UK). Paris: OECD.

OECD. 2008. *Sickness, Disability and Work: Breaking the Barriers*. Vol. 3 (Denmark, Finland, Ireland and the Netherlands). Paris: OECD.

OECD. 2009. *Sickness, Disability and Work: Breaking the Barriers*. (Sweden: Will the Recent Reforms Make It?). Paris: OECD.

OECD. 2010. *Sickness, Disability and Work: Breaking the Barriers*. (A Synthesis of Findings across OECD Countries). Paris: OECD

OECD disability web site. 2011. www.oecd.org/els/disability

OECD Social Expenditure Database—SOCX. 2008--www.oecd.org/els/social/expenditure.

Parsons, Donald. 1980. "The Decline in Male Labor Force Participation." *Journal of Political Economy*. 88(1): pp. 117-34.

Parsons, Donald. 1982. "The Male Labour Force Participation Decision: Health, Reported Health and Economic Incentives." *Economica* 49 (193). pp. 81-91.

Parsons, Donald. 1991a. "The Health and Earnings of Rejected Disability Insurance Applicants: Comment." *American Economic Review*. 81(5):1419-26.

Parsons, Donald. 1991b. "Self-Screening in Targeted Public Transfer Programs." *J.P.E.* 99(4): 859-876.

Parsons, Donald. 1996. "Imperfect 'Tagging' in Social Insurance Programs." *Journal of Public Economics* 62: 183-207.

Prinz, Christopher, ed. 2003. *European Disability Pension Policies*. European Centre, Vienna and Ashgate Publishing Ltd. England:

Rao, P. Madhava. 2009. "Social Security for Persons with Disabilities in India." www.dpiap.org/resources/pdf/social_security_india_09_04.07.pdf.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Belknap Press.

Riphahn, Regina. 1995. "Disability Retirement Among German Men in the 1980's." Discussion paper #95-20. Munich: University of Munich.

Rupp, Kalman and David Stapleton. 1995. "Determinants of the Growth in the Social Security Administration's Disability Programs—An Overview." *Social Security Bulletin*. 58(4): 45-70

Schimmel, Jody, David Stapleton and Jae Song. 2009. "How Common is 'Parking' Among Social Security Disability Insurance (SSID) Beneficiaries?" Ann Arbor, MI: MRRC WP2009-220.

Slade. F.P. (1984) "Older Men, Disability Insurance and Incentive to Work." *Industrial Relations*. 23(2): 260-277.

Soderberg, Elsy and Ulrika Mussener. 2008. "Entitlement to Sickness Benefits in Sweden: the Social Insurance Officers Experiences." *Environmental Health Insights*, 2: 13-23.

Stapleton, David and Kimberly Dietrich. 1995. "Longterm Trends and Cycles in Application and Award Growth." Paper presented at Conference on Social Security Administration's disability Programs. Washington DC.

Stapleton, David, Kevin Colman, Kimberly Dietrich and Gina Livermore. 1998. "Econometric analyses of DI and SSI Application and Award Growth." in Kalman Rupp and David Stapleton, eds. *Growth in Disability Benefits: Explanations and Policy Implications*. W.E. Upjohn Institute for Employment Research. Kalamazoo, MI.

Stapleton, David and Richard Burkhauser, eds. 2003. *The Decline in Employment of People with Disabilities: A Policy Puzzle*. Kalamazoo, MI: W.E Upjohn Institute for Employment Research.

Staubli, Stefan. 2009. "Tightening the Purse Strings: The Effect of Stricter DI Eligibility Criteria on Labor Supply." Netspar Discussion Paper 11/2009-041.

Waidman, Timothy, John Bound, and Austin Nichols. 2003. "Disability Benefits as Social Insurance: Tradeoffs between Screening Stringency and Benefit Generosity in Optimal Program Design." Ann Arbor: University of Michigan. MRRC working paper #2003-042.

Zaidi, Asghar and Tania Burchardt. 2005. "Comparing Incomes When Needs Differ: Equivalization for the Extra Costs of Disability in the UK." *Review of Income and Wealth*. 51(1): 89-114.

Endnotes

¹ For example, wheelchairs, hearing aids, caregivers' services, This has been estimated to add about 30% to the household's cost of living (Cullinan, Gannon and Lyons 2010; also see Zaidi and Burchardt 2005, Jones and O'Donnell 1995, Cutler, Landrum and Steward 2009).

² In a competitive insurance market, insurance companies will try to price according to the individual's risk category and an individual's decision about whether to purchase insurance implicitly reveals the value he places on the insurance. If he purchases, we know that the value of risk avoidance to him exceeds the insurance premium. The fact that many people do not buy disability insurance voluntarily, seeming to contradict the valuation by Chandra and Samwick, may be due to consumer myopia and underestimation of risk, or to the insurance company's inability to correctly differentiate risk categories.

However, in mandatory programs the individual's valuation of insurance isn't revealed, since they don't make voluntary choices. Also, strong limits are usually placed on differentiated pricing, so the system is redistributing as well as insuring and individuals will base their assessment on the redistributive effects as well as the pure insurance effects. Therefore, we don't have any "revealed information" about the pure insurance value of mandatory disability insurance. We only have simulated information based on assumed utility functions.

³ For the seminal work on tagging see Akerlof. For application to disability insurance see Parsons 1996, Diamond and Sheshinski 1995, Waidman, Bound and Nichols 2003.

⁴ Before 1984 82% of initial awards were based strictly on medical factors but by 2000 this had fallen to 58% at the initial stage and 40% by the time the final appeal had taken place (Autor and Duggan 2006).

⁵ An audit of Quebec's welfare programs in 1990-94 revealed mistakes in about 20% of the cases, most involving an overly-generous interpretation of which applicants met the means test. An audit in the UK in 1985 found both inclusion and exclusion error rates of almost 20% (Boadway, Marceau and Sato 1997).

⁶ Banks, Kapteyn, Smith and van Soest 2008 find wide differences across countries in the relationship between medical impairments and self-reported incapacity. In some high-income countries (e.g. the Nordic states), a large proportion of the elderly population consider themselves disabled, despite the high level of health and low mortality rates in these countries. In others (e.g. the Anglo-Saxon countries) the self-perceived threshold for considering oneself disabled is higher. Pain is the leading factor in reports of work incapacities and differences in reported incapacities are mainly due to different thresholds or response scales. Responses also vary by gender, education and age. This underscores the important role of social norms in determining what level of incapacity is unacceptable and worthy of social insurance support. Also see Kapteyn et al 2010.

⁷ However, disability is over-reported relative to objective measures among the unemployed, who are most likely to apply for disability benefits (Low and Pistaferri 2010).

⁸ One of the seminal papers on the theory of disability insurance, by Diamond and Sheshinski (1995), is based on a model in which disability benefits are greater than early retirement or

unemployment benefits and less than the full wage. Workers with a low disutility from labor work, those with moderate disutility apply for disability and work if it is not granted, and those with severe disutility (perhaps because of a medical problem) apply for disability and take early retirement or unemployment if it is not granted. The larger the disability benefit, the more people will apply and stop work. The impact of higher benefits on application, reciprocity and labor force participation rates keeps the optimal benefit size smaller than it would be otherwise. The trade-off between exclusion versus inclusion errors determines the optimal rigor of screening.

⁹ A large literature on worker's compensation claims also finds that incidence of muscular-skeletal and soft tissues claims are very responsive to stringency of the screening process and generosity of benefits. See Campolieti 2002.

¹⁰ Persons over age 55 previously faced a lax test for being disabled—they qualify if their ability to work is reduced compared with a healthy person with the same education in a similar occupation. In contrast, persons under age 55 only qualify if their medical impairment reduced their ability to work in a reasonable occupation for a person with his education. In 1996 this tighter criterion was applied to workers up to age 57; in other words, criteria were tightened for those age 55-57. Disability enrollments declined by 5.5 percentage points for those age 55-57, but not for other age groups.

¹¹ Also see Autor and Duggan (2006) and Duggan and Imberman 2008 on the positive correlation between disability reciprocity rates and unemployment rates over time.

¹² Golosov and Tsyvinsky 2006 propose screening via an asset test for eligibility for disability benefits: those who anticipate making a false claim will accumulate assets to smooth consumption over their lifetimes, so an asset test prevents false applicants from receiving benefits. (However, if potential applicants know that this test is being used, they may stop saving).

¹³ Numbers for Poland are sensitive to whether all disability schemes or only the main scheme are included. Using a narrower definition, OECD 2010 gives lower spending numbers than SOCX--2.8% of GDP in 1990 and 1.8% in 2007, compared with 2.3% and 1.9%, respectively, for the OECD average. Reciprocity rates are also lower in OECD 2010 than in OECD 2006. However the general direction of differences and changes is the same for all data sources.

¹⁴ The effective tax rate varies depending on hours worked. In the Netherlands, marginal tax rates are close to 100% for the partially disabled who work less than half their remaining capacity because their housing benefits and social assistance are phased out as earnings rise, leaving net income practically unchanged. Effective tax rates fall when the person works exactly half of remaining capacity, thereby qualifying for a wage supplement, then rise again as the supplement is unchanged while larger income taxes and social security contributions must be paid.

¹⁵ These individuals also retain their access to Medicare, publicly financed health insurance that is ordinarily available only to those over age 65. Disability benefit recipients under age 65 are eligible for Medicare. The value of this health insurance often far exceeds the value of cash benefits for disabled beneficiaries with high medical costs. If Medicare rights were lost, this would be an insuperable obstacle to their returning to work.

¹⁶ Many studies of wage subsidy schemes directed toward other disadvantaged groups (e.g. the long-term, low-wage or youth unemployed) come to conclusions that are consistent with these results. They show modest employment gains. Careful administration with tight targeting of benefits is necessary to ensure that they increase employment of the target group rather than subsidizing a broader group that would have worked anyway. But such targeting is difficult to achieve. These programs often involve considerable administrative cost and deadweight loss (Martin and Grubb 2001, Kangasharaju 2005, Lee 2005, Datta and Larsen 2008, Huttunen et al 2010).

¹⁷ Unless otherwise noted, the discussions of the Netherlands, Denmark and Australia in this section are based on data in OECD 2007, 2008 and 2010. Also see OECD 2003 and Prinz 2003.

¹⁸ An especially sharp drop in inflow to disability occurred in 2005, in part due to the lengthening of the sick leave period from one to two years. However, that short-run effect was largely dissipated by 2008 while the longer-term effects of the reform remained.

¹⁹ Several studies have examined the impact of experience-rated premiums in the unemployment insurance system in the U.S., using inter-state variations in these systems to identify impacts. They find large effects on lay-offs. Studies of the impact of experience rating and large deductibles on workmen's compensation claims in the U.S. generally find it is associated with a reduction in workplace claims and duration of work absences. This may occur because employers take greater care to avoid workplace injuries by investing in health and safety measures, because they use claims management to present countervailing information and speed the individual's return to work, or because their hiring practice becomes more selective. A study of experience-rating in Canada's workmen's compensation system found a decrease in incidence of short term disability claims but an increase in average cost per claim as more severe cases remained. Long term disability claims rose, perhaps for the same reason. Small firms are treated differently in this system because their accident rate may not reflect their true underlying risk. Another problem is that the true costs of some injured may not be known for many years (Hyatt and Thomason 1998).

²⁰ The shift to dual medical assessment of mental health claims also played a role--illustrating the importance of the screening procedure.

²¹ In Switzerland, too, employers have long provided disability benefits and sickness leave is widely covered by collective agreements rather than public arrangements. Typically employers contract with insurance companies to cover sickness and disability benefits and experience rating is used. We would expect the incentives facing employers and insurance companies to keep sickness leave and disability rates low. Indeed, Switzerland's disability inflow, reciprocity rates and sickness absences have long been among the lowest and employment rates the highest in

Europe. Employer provision has also created some of the problems observed in the Netherlands, such as the selection of healthy individuals in the recruitment process, which makes it difficult for older workers and other high-risk groups to change jobs, and competitive problems for small and medium-sized firms, whose incidence of illness has a large random component.

In addition, the Swiss system has a public program that is administered by the cantons but substantially financed by the central government—a mismatch between financial responsibility and decision-making authority, as in Denmark. Cantons have limited incentives to control costs and limited information—the individual’s own GP is the main source of information. Unlike the approach in the Netherlands, they have not used employers as a source of detailed information nor have they involved employers in developing rehabilitation plans. Switzerland also does not have the wage supplement that the Netherlands has introduced. So information flows and incentives facing workers, employers and public authorities are quite different in Switzerland. Their spending and benefit reciprocity rates have been rising due to population aging and a large influx of young adults with mental problems. (See OECD 2006 for more details).

²² For small firms experience rating would be based on the industry, not their own experience. Persons with severe clear-cut disabilities would be eligible for public insurance from the start.

²³ As in Australia, the UK uses private service providers, who are paid according to their success in placing disabled individuals in jobs. The UK’s Pathways to Work program does not change eligibility rules but requires new disability claimants to participate in 8 mandatory interviews; after that involvement with job counselors is voluntary. This program is aimed at getting the newly disabled back into the labor force quickly, rather than keeping them on benefits. (These are individuals with short run health problems who would be on sick leave rather than disability benefits in many other countries). Evaluations indicate that the outflow from the disability rolls was speeded up for individuals who would have left a short time later anyhow; and that women were more likely than previously to return to work (Adam, Bozio and Emmerson 2010; also see Bewley, Dorsett and Haile 2007, Lewis et al 2005).

As in most other countries, benefits are lost if the individual earns wages. To offset the consequent incentive problem, the UK offers a tax credit that serves as a wage subsidy for disabled workers in low-paid jobs, but this lasts for only 1 year. Since fear of losing the disability benefit appears to be a major reason why individuals do not want to leave beneficiary status, the UK allows the right to return if they find it difficult to work, within two years of the job start. But the take-up of these programs and the outflow from disability remains low (OECD 2007).

²⁴ Even though they could not exclude workers who wished to join a fund, they could make it difficult for those with a higher disability risk to enroll (e.g. by locating on the third floor of a building without elevators).

²⁵ Fees for disability and survivors insurance increased from less than 1% in 2008 to 1.8% in 2009. This may have been due in part to the loss of incentives for countervailing information. However, at the same time interest rates plummeted and rebates were required for women workers (because they have a lower risk of disability and death)—both of which would have contributed to the higher insurance premiums. Also, since 2008 the fee has been paid by the employer rather than the worker (although it may be passed on to workers in the form of lower wages).