

# **DISABILITY INSURANCE IN A MULTI-PILLAR FRAMEWORK**

Leo Aarts, Aarts & De Jong BV and Leiden University

Philip de Jong, Aarts & De Jong BV and University of Amsterdam

second version,  
November 1999

This paper was originally prepared for the World Bank conference: New ideas about old age security, Washington, D.C., September 14-15, 1999. The authors gratefully acknowledge helpful comments from Louise Fox and Jeff Hammer.

## **Abstract**

This paper describes the parameters of disability and disability insurance. Despite its complexity (work) disability is a privately insurable risk. But private insurance is at best incomplete, and leaves those most in need of protection uncovered. Governments fill this gap by setting up social insurance arrangements. Given the difficulty to accurately measure disability risks covered workers, employers, program administrators, or politicians may have interest in manipulating their occurrence. The element of choice in the use of disability insurance makes incentives the key concept to obtain a program design that is socially adequate as well as financially sustainable. Appropriate incentive structures can be found in systems that mix the redistributive advantages of risk pooling with the efficiency gains of competition. Such public-private mixes can be found in multi-pillar systems. The paper therefore deals extensively with disability policies in countries like Chile, Argentina, and Switzerland, and compares those with the experience of broad welfare states and transition economies.

## **Introduction**

Social disability insurance is designed to reduce the risk of earnings' loss due to impairments to function in gainful activity. The full private and social costs of disability exceed the losses related to reductions in productivity alone. They also include medical care expenses and the losses in social and psychological well being suffered by the person with impairment. But one person's impairment will also affect the well being of others. Family members and friends may have to make practical, financial, and psychological adjustments to accommodate people with handicaps. Society at large may also suffer some loss of well being simply by knowing of, or observing, the economic consequences of another's impairment. Such external effects are denoted as "collective compassion costs". They introduce feelings of solidarity as motives for the creation of private charities or collective government programs.

In the absence of any form of compensation, impairment costs would fall fully on people with disabilities, their relatives, friends, and other individual members of society. Social welfare, as conceived by economists to be the aggregate of each individual's welfare, would probably be lower than in societies that have some form of compensatory system to internalize collective compassion cost. Hence, even from the narrow economic efficiency perspective, a collective system can be justified by its contributions to increasing social welfare.

Compensation means transferring income from one party (disability program contributors) to another (disability benefit recipients). Social insurance is just one way to organize these income transfers. Private insurance and judicial liability assignments are conceivable alternatives to public arrangements. In practice, total compensation is often a combination of government benefits supplemented by a private pension. Such public-private mixtures are very common, although their relative importance may vary widely within and

across countries, depending on the industry or employment contract, and among programs by coverage and eligibility definitions.

Among a population that is covered by disability insurance being disabled and being a disability benefit recipient are two contingencies that do not fully overlap. Due to the elusiveness of disability these non-overlapping parts – disabled without benefits and non-disabled with benefits – will never be empty. Their sizes are an indication of the targeting performance of a disability insurance system.

### *Table 1*

As an illustration of the complexity of the concept of disability Table 1 shows prevalence rates of disability benefit recipients per 1,000 in the labor force. The countries in the Table are those that we use in this paper to compare national disability policies. Disability prevalence rates appear to vary widely across countries, even among similarly prosperous and healthy populations. The basic assumption in this paper is that such variation is largely due to the incentive structures induced by disability policy design.

The next two sections seek to give a conceptual basis for analyzing disability (insurance) policies. First we elaborate the concept of disability, followed by an analysis of the problems with unregulated market provision of disability insurance. The third section deals with two core parameters of general disability schemes. First, the variation in legally established disability insurance contracts defining coverage, eligibility, and entitlements is discussed. Then, the options regarding the public-private mix in the administration and funding of general disability insurance schemes are dealt with. The fourth section discusses work injury programs. The last section lists some recommendations.

This report is based on detailed studies of the (multi pillar) features of disability insurance schemes in Chile, Argentina, Poland, Latvia, Hungary, and Switzerland, more extensively. For each of these countries separate reports describing their disability arrangements are available from the authors.

### **The disability concept**

Disability is an ill-defined and complex phenomenon. It is not directly observable but must be inferred from its presumed causes and consequences. The World Health Organization (1980) defines disability as any restriction or lack (resulting from impairment) to perform an activity in the manner or within the range considered normal.

Assessment of impairments and their limiting consequences involves difficult and painful judgments by the individual afflicted and equally cumbersome verifications by the members of the individual's social setting. The inevitable subjectivity of these judgments makes individual tastes, social values, and financial evaluations part of the perception of disability. These subjective elements render disability a behavioral phenomenon. It is precisely for these behavioral aspects that disability, and disability insurance, are political issues, and fertile territories for social science and economic research.

In the definition of disability, health impairments are causally linked with reduced task performance. The presence of impairment is necessary but not sufficient. Some types of impairment have no incapacitating effects at all; others prevent performance but their adverse effects may be eliminated by training, medicine, aids, or appliances; still others imply total and irrevocable incapacity. Impairment is defined as an anatomical or psychological loss or other abnormality [Nagi (1969), WHO (1980)]. These losses and abnormalities may stem from diseases, traumatic injuries, congenital deformities, or prolonged disuse of muscles or

organs. They remain after the stage of active pathology has passed. The nature and severity of the underlying pathology determine the extent and permanency of impairment.

Impairments become manifest through the limitations in function or capacity they effect. Functional limitations can be grouped according to the level of organization at which they are defined and measured. One may discern limitations at the level of molecules, cells, tissues, organs, or the whole organism. Although limitations at a lower level of organization may not be reflected in higher levels, the reverse is not true [Howards et al. (1980)].

Evidently, limitations at the level of activities may result from diverse combinations of lower level limitations. For example, the inability to concentrate may be due to emotional factors, may have somatic origins, or both. And the inability to lift a heavy weight may be caused by energetical or locomotive limitations.

The obvious advantage of defining limitations at the level of activities is that they relate to the demands of task performance. Concerning work disability only those limitations are relevant that interfere with job demands. Thus, two people with exactly the same limitations may suffer different degrees of work disability. A hearing limitation is likely to affect a violinist more than a laborer, whilst a foot injury would affect the laborer more than the violinist. Impaired workers that are unable to meet the requirements of their old jobs may create alternatives to enrolment in disability benefit program through rehabilitation. Thus they may be able to acquire new skills and/or find a new job commensurate to their limitations.

### *Scheme 1*

Scheme 1 summarizes, from a clinical perspective, the causal chain leading from pathology to disability. According to this scheme, work disability can be defined as the inability to meet the demands of gainful activities, due to functional limitations, caused by

impairment. Two dimensions along which the elements of this causal chain should be conceptualized are severity and permanency. Other dimensions relate to the nature of impairment, limitations, and disability, e.g., a scale running from somatic to psychological, or refer to their cause: whether they are work-related or not.

### **Private Versus Social Disability Insurance**

In principle, private markets are able to provide disability insurance to a given population through the interaction of risk-averse consumers and profit seeking private insurance providers. Efficient contracts can be drafted such that policyholders maximize their expected well-being subject to the constraint that the insurers make normal profits. Such contracts provide coverage against income loss – the present value of foregone earnings during the period of disability. Each individual is charged a premium depending on his or her expected loss. Premiums, then, differ according to probabilities of disability, as indicated by age, health, and occupation. Per dollar covered workers with high disability probabilities are confronted with correspondingly high insurance premiums, and pre-existing disabilities are excluded from coverage.

Rating according to risk ('risk rating') is an inescapable consequence of competition. An insurance firm that would calculate an average premium rate so that low risk groups would subsidize high risk groups would run a loss because their low risk clients would be offered lower rates by competitors that charge rates equivalent to risk, and leave.

Since they value certainty risk-averse individuals are willing to pay an insurance premium that exceeds their expected losses. How much value they place on certainty depends on the depth of their aversion and on the size of expected loss. Using standard expected utility theory the value (V), or welfare surplus, derived from being covered can be expressed in money terms.

On the supply side, insurance policies will be offered only if the premium is high enough to cover the carrier's cost. Apart from transfers to policyholders who suffer losses the premium should also cover the transaction costs (T) involved in running the insurance business, for example the costs associated with the distribution of policies and the management of claims. Private insurance carriers will only supply insurance policies if the insurance premium at least covers the expected loss plus the transaction costs per policy.

An individual exposed to risk and an insurer will only engage in an insurance contract if the transaction cost is less than the money value of certainty induced by that contract (see Barr 1998, p.120).

$$T \leq V$$

Put differently, the social welfare gain from establishing an insurance contract equals  $V - T$ .

To the provider of disability insurance, whether private or social, the basic problem is defining and verifying the insured risk – income loss due to disability. Except in extreme cases, the severity of impairments, their incapacitating effects, their likelihood of cure, and the rehabilitative potential of those with lasting functional limitations, are difficult to assess accurately. Such assessment costs, and the other costs related to damage control, such as prescribing and monitoring preventive effort of the insured, and rehabilitative effort of benefit recipients, are all part of T.

Social adequacy, however, requires coverage of all members of a specified population (employees, self-employed, or the working age population) against the whole gamut of medical contingencies, and introduces a potentially sizeable amount of screening error into disability determinations (see Diamond and Sheshinski 1995; Parsons 1991). Two types of error can be distinguished: erroneous denials (exclusion or Type I errors) and erroneous admissions (inclusion or Type II errors). A first requirement to be met by a disability benefit

scheme is therefore that it should be designed such that it minimizes the sum of these two, mutually exclusive, types of error.

Associated with the delicate trade-off between exclusion and inclusion error are three insurance-specific problems that support the case for public provision of disability insurance contracts.

*Risk-dependence: unemployment and disability* First, individual disability risks may collectively depend on labor market conditions, thereby introducing *risk-dependence*. A private insurance market is only viable if the risk is shared among a heterogeneously composed population of policyholders. Private funds are likely to collapse under an unexpected surge of disability claims, provoked by an epidemic, high unemployment, or an increase in preferences for retirement.

To some extent disability and unemployment are linked. In a slack labor market more people with disabilities will try to seek shelter under a disability benefit program. Among a larger pool of applicants more people will be found eligible. Conversely, under favorable labor market conditions the number of people who are eligible for disability benefits but still work is larger than otherwise. This finding stresses that the decision to apply for benefits is often a matter of choice (see e.g. Rupp and Stapleton 1995).

Given the inevitable link between disability and unemployment, and the related problem of risk-dependence, disability schemes should exclude rules that allow consideration of employment opportunities in assessing the degree of disablement. In other words, in their assessment of disability program gatekeepers should as carefully as possible distinguish between the (medically based) inability to work and the (economically) based inability to *find* work that is commensurate with one's limitations. Gatekeepers, however, often find it difficult to disentangle the medical and the labor market factors that produce the disability claims that they have to adjudicate.

This administrative difficulty is aggravated by pressures – from politicians, labor unions, or individual claimants - to ignore this distinction and use disability schemes as an unemployment or early retirement program. Hiding part of a country's unemployment problem under the disability label may look attractive to politicians because lower official unemployment rates give a prettier picture of economic performance. For redundant workers being unemployed under the disability label may be both financially and psychologically more attractive as disability insurance schemes usually pay benefits of unlimited duration (until pensionable age) and are less stigmatizing.

Administrative practices by which the distinction between medical and economic disablement is blurred pose a major risk to the sustainability of social arrangements. Jobless workers that are recognized as disabled have an incentive to protect their livelihood by clinging to their disabilities and stay on disability as long as they can. They then lose their professional skills or their taste for formal employment. Being recognized as permanently disabled also means that workers are considered as, sometimes very early, pensioners who are not supposed to go back to work. Such practices create benefit liabilities that are larger and longer running than those for the regularly unemployed. Hiding cyclical, or structural, unemployment under disability creates problems that are more difficult to redress than the ones one seeks to solve.

A notorious example is the Netherlands that used its disability scheme as a generous and easily accessible exit route to counter the threat of massive unemployment in the 1970s. Between 1970 and 1980 the Dutch disability beneficiary volume more than doubled, and it kept growing until 1994, though at a slower pace. Although the program entry rate decreased in the 1980s, program exits stayed even lower. As a result, until today the Dutch social welfare system is kept hostage by a massive volume of disability beneficiaries. Similar problems plague East European transition countries (see Table 1).

*Lack of information: adverse selection and moral hazard* Second, those who provide disability insurance often lack reliable information on which to estimate individual risks and corresponding premiums, or even adequate grouping of the insured population into risk categories. If private insurance premiums do not closely reflect differences in expected losses, *adverse selection* can be a significant problem. Low-risk people will tend not to buy insurance, unless they are extremely risk-averse, whereas high-risk individuals will only buy insurance as long as they can afford it. The resulting high average risk among those who buy insurance will compel a correspondingly high premium, and participation in the insurance will further decline. Eventually the market will disappear unless insurers (are allowed to) invest in medical checks and exclude existing disabilities. Hence, due to adverse selection market supply of coverage will either be absent or exclude those most in need of coverage.

Third, the extent of risk may be influenced by one's willingness to work. By choosing the level of care taken to prevent or contain damages an individual can affect both the probability of incurring a loss and the amount of the loss. This is the problem of *moral hazard*. Moral hazard emerges if, upon full coverage, the insured acts as if damages would not involve costs and the returns to taking care were zero. This is again a problem of lacking information on the part of insurers, as they cannot predict the behavioral change among their insured populations.

The standard solution to correct careless risk behavior is to offer less than full coverage; or, in terms of social insurance against income loss, to offer replacement rates of less than 100 percent. The degree of coinsurance – the part of the loss that is left uncovered – is a matter of choice. “The degree to which it is desirable to reduce coverage and subject the insured to risk would depend on the incentive thereby created to exercise care, and such an incentive would in turn depend on the cost of taking care” (Shavell, 1979, p.541). The optimal

level of coverage is obtained when, for a given extent of risk aversion, the marginal cost of taking care equals the marginal benefit from reducing expected losses.

Since risk (experience) rating is a necessary condition to survive in a competitive setting market provision of disability insurance will help containing moral hazard as it rewards careful behavior through lower premium rates. Other ways to contain moral hazard are using strict eligibility rules and closely monitoring the care-taking behavior of the insured.

The presence of moral hazard reduces the welfare gain from insurance coverage, and therefore causes an efficiency problem. The size of the moral hazard component (MHC) under an insurance arrangement that covers income loss depends on the income elasticity of the demand for leisure. If leisure is a normal good higher replacement rates imply a higher MHC. The net welfare gain of insurance in the presence of moral hazard can simply be written as<sup>1</sup>

$$V - T - \text{MHC}.$$

This expression allows illustrating the trade-offs implied by the two instruments to contain moral hazard - co-insurance, and monitoring the risk behavior of the insured. Co-insurance means reducing both the value of certainty (V) and the moral hazard component (MHC). Gains are obtained as long as the reduction in moral hazard is larger than the loss of certainty. Similarly, monitoring increases T and reduces MHC.

The three insurance problems of risk dependence, moral hazard, and adverse selection are related to the inherent impossibility of accurate measurement of disability risks, and the possibility to manipulate their occurrence. It is for these behavioral elements, and the ensuing screening problems that disability risks differ from pension (longevity) risks. For these same reasons free market provision of disability insurance is unlikely to deliver an efficient

---

<sup>1</sup> See Aarts and De Jong (1998).

outcome. Both adequacy and efficiency require active involvement of the state, even more so than in the provision of old-age insurance.

### **General Disability Insurance**

When the market fails to provide adequate (universal, affordable) coverage of risks which have pervasive and drastic consequences people turn to government. At a minimum, government may respond by regulating the private insurance market, for example, by mandating purchase of a legally fixed standard insurance contract. However, as argued before, provision of social (mandatory) insurance by private competitors means risk rating and exclusion of people with existing disabilities. Additional programs, solidarity funds, or efficiency reducing forms of price regulation, then, have to take care of vulnerable groups that the market is unable to cover at affordable rates.

As an alternative a public monopoly that calculates pay-as-you-go contribution rates may offer base (first pillar) disability insurance covering every resident. This solution to the exclusion problem is found in almost all countries that have public disability insurance. However, as we will argue below, public administration of a pay-as-you-go fund may exacerbate the moral hazard problem (See also Lindbeck 1994). This, then, poses the basic dilemma in the design of social disability insurance: how to combine efficiency with adequacy?

Usually three separate programs cover the disability risk, each with its own legal base, funding, and administration:

1. A sickness benefit scheme covering short spells of (total) disablement of employees.

Often, these benefits are wage related and of limited duration with a maximum of one year. After lapse of entitlement to sick pay one may apply for a general disability benefit.

2. A work injury program covering work-related accidents and occupational diseases suffered by employees. Usually such a program covers both transitional and permanent disabilities. It distinguishes partial from total disablement.
3. A general disability program covering non-work related incapacity of employees or the population at large, including the self-employed, and those who are handicapped from birth or before they enter the labor market. Some general disability programs distinguish partial from full disablement.

In countries where this set of programs exists they usually have a historical background. But this separation may also serve to make the coverage of disability risks more manageable. Risks associated with short-term and work-related contingencies may be easier to calculate. Sick pay coverage usually has a limited entitlement duration whereas work injury risks are more narrowly defined.

This is illustrated by the fact that in some countries sick pay and work injury liabilities are privately covered by mandating employers to pay for these benefits (U.S. (sick pay and workers compensation), U.K. (sick pay), Netherlands (sick pay), Finland (work injuries), Chile (work injuries)). Treating these “easier” risks separately allows policy makers to focus on the program that is the more difficult to manage – the base general disability insurance scheme.

The variety in disability beneficiary volumes across countries (Table 1) can be largely attributed to the position chosen in reaching the conflicting aims of adequacy and efficiency. This position is best described by the parameters of the programs that cover disability risks. In the remainder of this section we elaborate two of these parameters for a general disability scheme. The first is the legal framework: the rules that determine coverage, eligibility, and entitlements. The second is the public-private mix in administration and financing.

### *The Legal Framework*

Two basic features of social insurance are (1) mandatory consumption of (2) a legally established insurance contract (see Scheme 2 below). Mandatory participation in an insurance arrangement solves the adverse selection problem because low risks cannot opt out. It also solves the free rider problem, although contribution evasion may be a serious problem in societies with low tax morale.

*Coverage* If a country seeks to protect its citizens against the financial consequences of disablement the coverage conditions should be broadly defined. Depending on political preferences, and a country's prosperity, coverage may include all citizens, or only the labor force. For instance, broad and prosperous welfare states such as Norway, Sweden, and Switzerland, have one national (first pillar) program that covers the entire working age population. On the other hand, middle-income countries like Argentina and Chile only cover that part of the work force that, through its employment status, is an obligatory participant in the social security system. Intermediate cases are countries like Germany, and Austria, which cover the complete work force but have no programs for those without an employment record. Finally, countries such as the United States and the Netherlands have wage-related programs for employees, and a separate (flat benefit) program for those without a (sufficient) earnings record. Often, full coverage is only obtained after a certain number of "insurance years". With few exceptions, coverage ends at the age when coverage under a national pension program starts.

Coverage not only refers to the insured populations but also includes the part of earnings that is covered by social insurance. Almost every public disability insurance program has an income ceiling above which earnings are not covered. These ceilings are usually between one and two times the average wage. This is another type of coinsurance. Those with

earnings beyond the ceiling can buy additional coverage at the private insurance market. The same applies to groups, such as self-employed, that are entitled to flat rate benefits only.

*Eligibility* Because disability insurance programs aim at those with a long-term, or permanent, inability to work, eligibility generally starts after a mandatory waiting period during which claimants must have been without earnings as proof of their disability. Many programs have a waiting period of one year, covered by Sickness Benefits. For lack of a federal sick pay scheme the U.S. Social Security Disability Insurance has a waiting period of only five months.

Similarly to coverage, eligibility for disability benefits ends when eligibility for national pensions starts. Only in very few countries – Poland is an example – people can choose between disability benefit or an old-age pension. Since for many people in Poland disability benefits are higher than social security pensions the number of disability beneficiaries includes a sizeable number of pensioners (see Table 1).

Given the elusiveness of disability, and the ensuing problems of risk-dependence and moral hazard, countries that define the disability risk in stricter, less ambiguous, terms have programs that are more manageable. In line with Scheme 1, risk definitions run from a medical assessment of impairment, or functional limitations, to a measure of loss of work, or earning, capacity, which involves a confrontation of limitations with job demands. Work capacity definitions add difficult to measure vocational aspects to the, already cumbersome, determination of functional limitation.

The Dutch disability insurance scheme gives an example of an elaborate disability definition. It measures the extent of disablement by considering the residual earning capacity. This capacity is defined by the earnings flowing from any job commensurate with one's residual capabilities as a percentage of pre-disability earnings. The degree of disablement, then, is the complement of the residual earning capacity. Such a definition allows for a

continuous scale, on which to measure earnings loss, and a corresponding scale of compensation. The Dutch approximate this continuous measure by distinguishing seven disability categories ranging from 15-25 percent, to 25-50 percent, and so on, up to 80-100 percent disabled. Replacement rates correspond to these categories but never exceed 70 percent of covered earnings. Other countries with similarly complex definitions are Germany, Switzerland, and Norway, but these countries do not have as many (partial) disability categories.

In stead of an earnings capacity definition most countries use loss of work capacity as the risk covered under disability insurance. Such a definition confronts functional limitations with activities that are normally required in paid work. The extent to which activity performance of an impaired worker is below what is normally required defines the percentage reduction in work capacity. These assessments do include vocational effects of impairment but stay short of measuring the effect of reduced capacity on earnings. As the examples of such diverse countries as Sweden, Poland, and Chile, show work capacity ratings do not preclude partial disability categories. For instance, Sweden has four disability classes (starting from 25 percent disablement).

Two more aspects may further complicate disability definitions. First, schemes differ in the type of jobs they consider suitable if one loses the capacity to perform one's current, or usual, line of work. Suitable jobs are those that are compatible with one's training and work experience. Under such ruling a worker is considered fully disabled if he is unable to do suitable work. Second, some systems that recognize partial disablement award full benefits to those partially disabled who are unable to find a job that is commensurate to their residual capacity. Sweden, and Germany used to have stipulations that allowed workers over 60 years with sufficient insurance years to enter disability rolls for labor market reasons, without a medical examination. Both countries abolished this rule. Until 1987, the Netherlands

explicitly recognized the difficulties partially disabled workers may encounter when looking for work. This rule was applied so leniently that almost 90 percent of new awards were based on full disablement irrespective of the extent of the underlying medical disability. Such labor market considerations blur the distinction between unemployment and aggravate the moral hazard problem considerably.

Summing up, the inherent slack in the definition of disability risk may be reduced by using unambiguous assessment protocols. Disregarding the vocational consequences of impairment, and defining the risk in terms of health contingencies, may make eligibility checks easier but it targets the unhealthy and not the disabled.

Second, the set of commensurate jobs by which residual earning capacity is measured is crucial, too. Accessibility to compensation can be reduced by using a broad set, like any gainful activity, in stead of relating functional limitations only to jobs that are suitable with respect to one's education and work history. Third, labor market consideration (availability of commensurate work) should be excluded from disability determinations. If one allows for such consideration a disability benefit scheme will cover unemployment risks as well, and will become even more difficult to manage.

And, finally, the question remains whether partial benefit categories are useful. Partial benefits may work as a wage subsidy by allowing handicapped workers to offer themselves at a reduced price in order to compensate employers for their lower productivity. In actual practice, however, disability benefit adjudicators often find it difficult to make assessments within the discrete disability categories that are laid down in law, and tend to award full benefits to the partially disabled. In that case they widen the gates to disability benefits and are a disincentive instead of an incentive to return to work. In countries where benefit levels are closely linked to the degree of disability, as in the Netherlands and Sweden, most awards

are based on full disablement. At the end of 1995, 75 percent of Dutch disability beneficiaries, and 85 percent of the Swedish disabled had an award based on full disability.

*Entitlements* Disability insurance entitles primarily to cash benefits. Some programs also include benefits in kind, such as vocational rehabilitation, wage subsidies for employers, and subsidies for work place adaptation. Often, rehabilitation measures are covered by a separate program, or by health insurance.

Generally speaking there are two types of cash benefit systems for general, long-term or permanent, disability. The first is truly multi-pillar because it consists of two tiers. Sweden and Switzerland are examples of such systems. Each working age resident that meets the eligibility criteria is entitled to a flat benefit. Disabled employees are entitled to a second tier benefit under the same eligibility standards. This second tier benefit supplements the base benefit up to a percentage of covered earnings. As mentioned, private insurance may cover earnings surpassing the social insurance coverage ceiling. The second type of benefit system has no universal base benefits but awards earnings related benefits to employees, and has separate schemes for special groups, like the self-employed and the youth handicapped, awarding a flat benefit.

Wage-related benefits replace only part of earnings lost, even when fully disabled. In many countries (Argentina, Chile, Latvia, United States) statutory replacement rates depend on contribution years. Minimum income guarantees usually protect disabled workers, and their households, from destitution. Income floors imply that low-income groups have higher replacement rates than those with higher incomes. In 1993, average before tax replacement rates under general disability insurance programs varied from 30 percent in the United States to 74 percent in Sweden (Aarts et al., 1998, Blöndal and Pearson (1995)). In European welfare states statutory benefits are often supplemented by collectively bargained

supplements, leading to considerably higher effective replacement rates, and correspondingly weaker work incentives.

To prevent workers with disabilities from leaving employment the United Kingdom and Latvia provide working allowances. These allowances are similar to partial benefits in that they seek to compensate employers for the lower productivity of impaired workers. But they have two major advantages: first, while partial benefits can only be awarded upon entry onto the disability rolls working allowances can be obtained without leaving the job. Second, working allowances are easier to administer and not subject to the screening errors that may lead to disincentive effects of partial benefit systems.

## *Scheme 2*

### *The public-private mix*

In Scheme 2 we give an overview of differences between public and private insurance. Previously we argued that coercive consumption of a legally established insurance contract is the prime characteristic of social insurance. A second crucial characteristic of social insurance is risk sharing. The way in which premium rates are differentiated among the insured population reflects the degree of (risk) solidarity under an insurance contract. Degrees of risk sharing run from a nationally uniform percentage of earnings to rates which vary by individual risk profiles. However, individual risk rating is incompatible with mandatory participation because high-risk persons would face a forbiddingly high premium rate. Therefore, mandatory insurance always requires some form of pooling through which low-risk groups (are forced to) subsidize high-risk groups.

Often, disability risks are pooled nationally, and financed by pay-as-you-go. Pay-as-you-go financing can only be used to cover large financial liabilities – such as long term

payment of disability benefits – if the insurance has a broad base, and the mandatorily covered population cannot choose among insurance funds. In other words, pay-as-you-go is only feasible if a public monopoly fund acts as the social insurer. In the absence of moral hazard the cost of providing social insurance through a public monopoly is potentially considerably lower than market production. The product is a non-negotiable standard contract, purchase is mandatory which eliminates marketing expenses, and there are large economies of scale.

In the presence of moral hazard, however, the problem with pooling is that the higher the level of risk sharing the smaller the returns to avoiding, or limiting, damages. When a disability program is financed by nationally uniform pay-as-you-go rates the cost of inappropriate use is shifted to a national fund. Micro economic incentives to control losses, then, are eliminated, and the only remaining price that signals abuse is the aggregate contribution rate.

This incentive (moral hazard) problem may not only affect covered employees, but also their employers and the program gatekeepers. Employers may prefer disability insurance as an instrument of employment policy if the disability route is cheaper and administratively less burdensome, than openly terminating an employment contract. Likewise, program gatekeepers who operate in a public monopoly setting may be inclined to pursue a conflict avoiding strategy towards claimants, both to reduce their workload and the psychological burden of being strict. Such attitudes lead to lenient interpretation of eligibility rules, even when they look stringent on paper.

The issue of careful application of eligibility rules points to a crucial role of those responsible for claim management in limiting the moral hazard problem. One of the core problems of public management of pay-as-you-go disability insurance programs is the lack of cost consciousness among those who screen eligibility conditions as they are not confronted

with the financial consequences of their adjudicative practices. Examples of administrative slackness can be found under a variety of disability schemes like those in the Netherlands, Poland, and the United Kingdom. Other countries, like Germany and the United States, successfully use bureaucratic control mechanisms as a substitute for financial discipline.

There are three ways in which the incentive problems of insurance and administration of disability risk by a public monopoly may be reduced. The first is found in some of the countries where disability benefits are an integral part of the public pension system. In countries like Germany, Austria, Switzerland, and Sweden, disability benefits are not awarded as long as there is any rehabilitative potential left. By giving priority to rehabilitation over benefit dependency these countries try to target pensions to those workers who are fully, and irrevocably, incapacitated.

In these countries rehabilitation is mandatory for workers with disabilities and this mandate is supported by a comprehensive infrastructure of rehabilitation services. In Germany, a special public agency assigns severely handicapped statuses to workers that have lost 50 percent or more of their work capacity. This agency, which operates independently from the pension funds, focuses on keeping impaired workers in the labor force instead of awarding benefits. Workers with a severely disabled status have certain privileges, and help their employers to meet the statutory quota obligation of employing one severely disabled person for every 16 job slots, or pay a fine. This agency also gives subsidies to accommodate jobs. Germany and Sweden, however, treat workers over 60 more leniently, and exclude them from mandatory participation in a rehabilitation program.

A second way to improve incentives is by outsourcing administrative tasks to private parties under public monopoly insurance. The Argentine work injury program exemplifies such a mixed model. In Argentina employers are obliged to cover their workers against work-related disabilities by contracting with an Aseguradora de Los Riesgos de Trabajo (ART).

ARTs collect premiums, pay out benefits, implement prevention programs, and organize rehabilitation. Contrary to their name, ARTs do not carry the insurance risk. A public monopoly insurance fund calculates contribution rates, which are differentiated according to the risk profile of a firm. Firms, therefore, have an incentive to choose an ART on the basis of its record in containing damages.

It is somewhat surprising that Argentina uses a mixed model to cover work-related disabilities which are privately borne and administered in several countries, while it is among the very few countries that have opted for full privatization of both the administration and the coverage of the general disability risk. Using a model that mixes two allocation mechanisms – public monopoly insurance and privately organized administration and rehabilitation – prohibits reaping the full efficiency gains of a competitive insurance market. More efficient results can be obtained if one allocation mechanism governs both insurance and administration. In that case investments in monitoring beneficiaries can be efficiently weighed against reductions in benefit expenditures.

*Private delivery of the full insurance function*      A third alternative to discipline the parties involved in disability insurance is delivery of the full insurance function (administration and funding) by private competitors. Choosing a lower pooling base – private or self-insurance - may solve the incentive problem of public monopoly insurance, and the lower the pooling base the stronger the incentive to use disability insurance carefully. Despite the enormous problems in managing general disability schemes, however, very few countries use lower bases than national funds. Chile, Argentina, and Switzerland are among these few countries. Chile and Argentina use private pension funds to pool disability risks, and Switzerland covers its second tier disability liabilities through firms.

Mandating firms to carry the disability risk of their personnel is technically possible, and socially defensible. As described in Scheme 1 disability is the result of functional limitations that prevent workers from meeting job demands. By adapting job demands, and work conditions, employers may influence both the probability of their occurrence and the duration of disablement. Disability, therefore, is an employment risk, and employers can be reasonably asked to carry it.

An important advantage is that self-insurance confronts firms with the full insurance liability. They have a strong incentive to reduce moral hazard to its efficient level, where the marginal benefit of reducing inappropriate use is equal to the marginal cost of monitoring the risk behavior of their employees. Within such an incentive structure employers also efficiently weigh investments in prevention and job accommodation against lower benefit expenses. Moreover, organizing the coverage of workers' risks at the firm level offers the possibility to take account of specific preferences of a firm's personnel.

Whether the firm self insures, or has its disability liability reinsured by a private carrier, funding of its benefit payment obligations is necessary as the disability risk may involve long-term payments. Usually private insurers finance disability risks by *pre-funding*. Pre-funding means calculating an annual premium sum such that it covers the present value of the benefits that are expected to start in the current year. Pre-funding, therefore, is something between pay-as-you-go financing and full capital funding. It is neither entirely neutral with respect to aging of the population at risk, as aging increases the cohorts entering the disability insurance scheme, nor is it inflation proof. On the other hand, if disability entitlement stops at pensionable age and the average age of those who become disabled increases the average duration of benefit payment will decrease, and so will the pre-funding premium rate. Moreover, although the funds created by pre-funding are much smaller than the savings necessary to finance pensions, they still yield additional sources from investment returns.

Another major advantage of pre-funding is that it does not entail the portability problems that plague capital funding. Employees can move freely from one firm to the next without jeopardizing their disability benefit entitlements. Equally, firms who reinsure their benefit liabilities may change insurers at low cost.

Making firms responsible for the payment of social (mandatory) disability benefits also has a number of drawbacks. The potential efficiency gain from employer based delivery of social insurance comes from the potential conflict of interest between employers, who want to keep their non-wage costs low, and their employees who want adequate protection against wage-loss risks. Firstly, firms may be overly strict and deny benefits to claimants even if they meet the eligibility conditions, to keep benefit expenditures, or experience rated premiums, low. Similarly, under pressure of the risk of future benefit payment obligations firms may be unwilling to employ persons with a bad health record.

Most countries have legal systems that allow workers to litigate against such adverse decisions. Litigation costs are an inevitable byproduct of private delivery of social insurance. Secondly, employers may find it unacceptable that the lawmaker holds them liable for the payment of disability benefits in cases where the cause of disablement entirely lies in the private sphere (skiing accidents, stress from divorce, et cetera), or where health complaints are not objectively assessable. Thirdly, small business has an equally small pooling base, and may therefore face high insurance costs. Organizing a broader base through small business associations can solve this problem. A final problem with private delivery of social disability insurance is that it excludes persons who became disabled before entering the labor force. As mentioned, these youth handicapped may be covered by a first pillar arrangement (Sweden, Norway, Switzerland), or by a separate scheme (Netherlands, U.S.), or by a special provision under the program covering Child Allowances (Argentina, Chile).

*Multi-pillar disability insurance* As an alternative one may pool general disability risks through private pension funds. This is done in countries like Chile and Argentina, but also in several European countries where supplementary benefits are provided by occupational pension funds. However, for three reasons the Argentine and Chilean arrangement is a unique case of private social insurance. First, up to an earnings limit these arrangements cover the full disability risk – not just a supplement. Groups that do not take part of this employee based (second pillar) system, and workers with small pension savings, are covered by a minimum benefit provision financed through pay-as-you-go contributions.

Second, the pooling base is a private pension fund chosen by the individual employee, and each employee of a firm may select a different fund. These defined contribution (“individual capitalization”) funds are obliged to accept all workers as “affiliates”. Apart from managing pension accounts they also contract disability insurance for their pool of affiliates with a private insurer. These group policies cover temporary disability benefits and, in case of permanent disability, they transfer an additional lump sum to the pension account so that the pension fund is able to pay an annuity over the rest of the disabled worker’s life. The insurer is obliged to calculate average (group) rates, creating solidarity between low and high-risk groups.

And third, disability assessments are not part of the privatized insurance and administration services. Independent medical committees check whether a claimant meets the eligibility requirements. Such independent committees solve the potential conflict of interest between labor and management. As mentioned such conflicts may arise when private insurers falsely deny benefits to keep expenditures low.

Using a multi pillar framework for disability insurance that is part of, and therefore consistent with, the pension scheme has several advantages. First, it has the same advantages of multi pillar pension schemes, such as combining redistribution (income solidarity), and

inflation proof benefits, in the first pillar with individual responsibility, and demographic neutrality, in the second pillar (World Bank 1994).

Second, by making disability insurance part of the pension system the emphasis is on rehabilitation, and permanent pensions are only awarded to those without prospects of recovery. Note that this does not rule out partial benefits. For instance, since 1992 Chile has four disability categories (15 - 40 percent, 40 – 70 percent, 70 – 100 percent, and 100 percent disabled). Temporary benefits are used to cover the period up to when the state of maximum medical improvement has been reached.

Third, having disability and old-age pensions financed and administered by the same funds (AFPs) allows for efficient insurance of permanent disability benefits. Permanent disability benefits are funded by a pension (savings) account. If the accumulated capital is too small to pay out the statutory disability benefit the insurance company adds a lump sum. The fact that disability risks increase with age but the additional sum needed to fund disability benefits up to pension age decreases with age makes for a limited and calculable risk. Mature private insurance markets have no problems covering such risks. This allows profiting from the efficiency gains engendered by delivery of social insurance through competition, and from economies of scope induced by combining two social insurance arrangements.

The Swiss disability arrangement is similar to that in Chile and Argentina and has the same advantages. It has also the same structure, and is part of, the multi pillar pension scheme. It differs in the sense that it is more truly multi pillar. Base benefits for employees are covered by the universal disability program (first pillar) which insures the entire population. On top of this first tier benefit - the size of which depends on contribution years – employees are mandatorily insured for a supplement. These second tier benefits are funded by the same combination of savings and insurance as the Chilean arrangement described above. The insurance part consists of a lump sum that adds the statutory contributions to the pension

account for years between the age at onset of disablement and pension age. Interest payments that would have accrued on the pension account if a worker had stayed on the job are disregarded in calculating the insurance sum.

### **Work Injury Programs**

Most social welfare systems started by covering work-related disability risks. The 19th century Industrial Revolution confronted countries with new risks. Workers were injured, or even killed, by industrial machinery that was new and unfamiliar. Labor unions started developing funds that offered their members coverage of the financial consequences of work injury risks. These mutual funds were organized by various industries, and coverage differed accordingly. The extent of coverage depended on the risk, and on the coercive power of the labor movement in the relevant industry, but usually included wage loss, health care, and survivor benefits. Contributions were levied on a pay-as-you-go basis and participation was mandatory.

Injured workers that were not covered by these union funds or by funds that compassionate employers had established could only appeal to the liability of the employer, and try to get compensation through the courts. Judges increasingly acknowledged employers' liability in causing injuries, and forced compensatory payments. Hence, what started as a life-threatening risk for workers developed into a financial liability for employers? This created a broad political platform, including both labor market parties, for social insurance of work-related risks. Both parties agreed on a *quid pro quo* implied by legally established policy conditions: by accepting the mandate to pay a legally defined compensation employers were protected against unlimited tort claims, and by accepting a defined compensation workers were protected and released from the obligation to prove fault. Hence, in all industrialized

countries government started to intervene in disability coverage with the enactment of an Industrial Accidents Act.

Previously we argued that mandatory insurance against general disability risks might be provided by private insurance instead of the public monopolies that are found almost everywhere as the providers of disability insurance. If one is willing to accept the validity of this argument it certainly applies to the case of work-related contingencies as they are a narrower, more strictly defined, class than general disabilities and expected damages can be predicted more reliably. For that reason social insurance against work-related disabilities is covered by private insurance in high-income countries like the United States, Switzerland, Finland, and Belgium, but also in a middle income country like Chile. Be it that these markets are heavily regulated so as to reach an acceptable balance between efficiency and accessibility.

Across countries Work Injury programs are even more heterogeneous than general disability programs. Generally speaking their design is more like private insurance arrangements than that of general disability programs that often have more of a redistributive aim. Another common trait across different programs is that employers pay contributions. But, as economic studies have shown, firms may shift such non-wage costs to employees by lowering wages (Fishback and Kantor 1995). Other distinguishing characteristics of Work Injury programs can be ordered under the same two headings as those that we used to describe general disability programs: legal framework and public-private mix.

### *The Legal Framework*

By definition, Work Injury programs are more restrictive than their general counterparts, both in terms of coverage and eligibility. Coverage is limited to work-related injuries and occupational diseases but the dividing line between work-related health contingencies and

others is in some countries rather thin. Employees are the prime target group for coverage but students and trainees may also be covered, while self-employed often can participate on a voluntary basis. From the start of their employment contract workers are fully covered.

Switzerland and the Netherlands are special cases. Switzerland mandates all of its residents to take out accident insurance. Swiss employees and unemployment beneficiaries are insured under the Accident Insurance Law covering both income loss and health care due to accidents irrespective of cause. The damages under this law are split into work-related accidents and others. The work-related part of the insurance also covers occupational diseases, and is paid for by charging the employer. For non-work related accidents a separate premium rate is calculated and paid by the employee.

The Netherlands is the only country that does not distinguish between general and work-related disablement. In 1967 it introduced a comprehensive Disability Insurance scheme that covers every health contingency that reduces one's earning capacity, *irrespective of cause*. The design of this program derives more from the previously existing Work Injury program than from the considerable less generous general disability program that prevailed before.

The differences between general and work-related disability programs are related to generally accepted rules of tort law: to the extent that they can influence occupational risks to which their manpower is exposed employers are held responsible for the prevention and indemnity of these risks. This rule is translated in economic terms by means of financing schemes that are based on the polluter-pays-principle: firms that invest in controlling disability damages are rewarded by lower premium rates, and vice versa. With very few exceptions, therefore, Work Injury programs that are run by a public monopoly insurer calculate premium rates that differ according to industry, or according to the individual firm.

The eligibility rules of Work injury programs are usually more transparent and explicit than those covered by general programs. Often covered risks are limited by using an exhaustive list of (causes of) injuries and diseases. Such transparency serves three purposes:

- (1) it limits employers' liability and, therefore, broadens political support for a program that mandates business to provide coverage;
- (2) it makes risks more easily calculable, facilitates private insurance supply of coverage and, therefore, promotes efficiency;
- (3) it eases the distinction between the risks covered under a Work Injury program from those covered under general disability and, therefore, reduces the risks of cost shifting by firms and of selective behavior by workers who are covered by both programs but prefer Work Injury benefits as they often are higher.

This last point refers to the way in which entitlements under Work Injury programs are assessed. As always they depend on the severity and expected duration of the functional limitations caused by occupational hazards. As a general rule programs distinguish between temporary and permanent disabilities, and between partial and total disablement. Like under general disability programs severity is measured by choosing a point on the causal chain running from impairment to inability to perform gainful activity (Scheme 1). Under the first Industrial Injury Acts the typical risks covered were traumas affecting specific parts of the body, such as the loss of a limb. Elaborate schedules (point systems) were developed to express such losses in monetary terms. For example, loss of a finger would entitle to 10 percent of covered earnings, and loss of a hand to 33 percent.

Such (old fashioned) compensation of losses of scheduled body parts still exists. Often, these impairment-based entitlements are not earnings tested. They can be awarded as annuities or as a lump sum. Because they do not take account of the requirements in one's usual work these impairment ratings do not measure loss of work capacity, and are therefore

not recommendable as risk definitions under a social insurance scheme that seeks to compensate the vocational consequence of injuries.

It is more preferable to consider the impact that occupational hazards may have on a worker's ability to earn or to compete in the labor market, and to use assessment procedures that are equivalent to those that are used in general disability programs. Work Injury programs, however, have ratings that are more refined, and often use an extensive set of partial disability categories. Among the middle-income countries studied for this report Chile and Argentina have four categories, and Latvia eleven.

Evidently, under a loss of work capacity definition benefits are earnings tested. As a rule they are higher than general disability benefits which causes pressure from impaired workers who want to be recognized as work injured instead of regularly disabled. For instance, in Latvia total incapacity entitles to 80 percent of insured earnings if it is work-related, and at most 55 percent otherwise. And in a multi-pillar country, like Chile, work injury benefits have a fixed replacement rate (70 percent of earnings when totally disabled) whereas general disability payments depend on savings and insurance policies with only a minimum income guarantee.

Another distinguishing characteristic is the fact that medical and vocational rehabilitation services, and physical in kind provisions, are integral parts of Work Injury programs while general programs often lack such instruments to contain beneficiary volumes. Moreover, as a rule, participation in rehabilitation is mandatory. The stronger emphasis on rehabilitation under Work Injury programs is another reflection of the fact that occupational hazards are seen as a risk of the firm. Firms are responsible for a safe work environment and have to make room for workers with residual limitations after the onset of work-related injury or disease.

### *The public-private mix*

Following the efficiency rule that risks should be pooled at the level where they can be best influenced Work Injury programs use premium rates that are differentiated by risk. Countries that recently introduced such schemes like Poland, Latvia, or Hungary, started with a nationally uniform pay-as-you-go rate. From the year 2000, Poland will calculate contributions that vary according to industry and range from 0.40% to 8.12% of covered earnings. Gradually the number of risk factors in the calculation of premium rates will be increased, and eventually every firm will be charged according to the risk profile of their personnel and production process.

In financial terms it is not such big step to move from actuarially equivalent rating under a public monopoly to allowing firms to self-insure, or to buy Work Injury insurance in a competitive market. Each of these options is preferable over pooling risks at a level higher than the individual firm. Nationally, or industry-wide, rating forces inter-, or intra-, industry solidarity at the expense of eliminating the financial returns to prevention and rehabilitation that risk rating yields.

Although they are financially equivalent the two remaining options - public monopoly provision of insurance at actuarial rates and private supply of a legally fixed package of coverage – are different in other respects. Public monopoly supply entails public administration (claims management). Many European countries (France, Germany, U.K., Sweden, and others) have such a public benefit allocation structure. Although it not a big issue in these countries, public monopoly allocation may cause a conflict of interest between firms that want to contain their contributions and public adjudicators that are biased towards the injured worker. Unambiguous risk definition, transparent assessment procedures, independent medical committees that do the assessments, and an accessible and efficient appeals procedure, can reduce the risk of conflict under a public regime.

Private market supply of the full insurance function – financial (actuarial) as well as claim management – as the experience with the U.S. Workers Compensation program shows induces another risk. In most American states firms may choose among private insurance carriers to have their Workers Compensation liabilities covered. For profit maximizing reasons firms and insurers may then collide to keep damages low by being excessively strict. This has been an issue in a number of jurisdictions. This risk may be counterbalanced by the same set of instruments that can be used to offset the bias under a public regime.

Interestingly, in Chile and Argentina the same independent medical committees that also do general disability assessments adjudicate work injury claims. In Germany, Industry Associations that have a legally protected sectoral monopoly run the Work Injury program. Every employer is legally mandated to buy Work Injury insurance from the Association of the industry in which he operates. Firms are charged actuarial equivalent rates. Each Association contracts, or owns, and a broad set of rehabilitation services. Finally, Switzerland has its (work-related) accident insurance supplied by 41 insurance companies of which 29 are private competitors. It also has a very limited list of recognized (causes of) injuries and occupational diseases. All other disabilities are covered by general disability insurance.

*Vulnerable groups* One major drawback of pooling disability risks at the level of the firm is that they may discriminate against workers with handicaps. Workers with a preexisting health problem pose a financial risk to employers that are legally obliged to insure wage loss due to accident or disease. Under most of the U.S. Workers Compensation statutes this problem is solved by creating a special (solidarity) fund for workers with preexisting conditions. These so-called Secondary Injury Funds protect employers from being obliged to pay for contingencies that may be related to previous work but not to the current employment conditions.

Discrimination against handicapped workers for financial reasons is much less of a problem if risks are pooled in pension funds, as is the case in Chile and Argentina. The major advantage of these funds is that they primarily accumulate capital. Their primary task is to provide for longevity risks, and disability has only a minor part in their risk portfolio. In fact, to some extent these risks hedge each other: disability and longevity risks are negatively correlated.

### **Recommendations**

Social disability insurance is designed to reduce the risk of earnings' loss due to impairments to function in gainful activity. This risk is multidimensional and complex. First of all, disability is not directly observable but must be inferred from its presumed causes and consequences. Second it has both medical and vocational dimensions. Work disability can be defined as the inability to meet the demands of gainful activities, due to functional limitations, caused by impairment.

Despite its complexity disability is a privately insurable risk. To make the disability risk a predictable one private insurers use restrictions. First they exclude pre-existing health conditions from coverage; and if these conditions are serious enough to prevent substantial gainful activity they deny coverage. Those handicapped from birth, or in their youth, are likely to be denied insurance. Second, private insurers tend to define disability in terms of impairments because they are easier to measure. As a consequence, they ignore the vocational consequences of impairment.

In the absence of social disability insurance the cost of providing for these excluded groups is shifted to other social welfare programs – such as public assistance – to their family, or to other caretakers. To counter this market failure governments have set up social insurance

arrangements. The principal difference between social and private disability insurance is that a social arrangement implies mandatory purchase of a legally fixed insurance policy.

Moreover, social insurance schemes offer broad coverage and do not use exclusions. But if disability is defined in broad terms, accurate assessment of the extent of disablement becomes an even more complicated task. The broader the definition of disability the larger the discretion of the program gatekeepers, but also claimants and firms, to use disability benefits according to their own preferences. It is crucial that these interest parties (including governments) are confronted with the financial consequences of their choices.

There are two broad sets of instruments that can promote efficient use of disability insurance. First, the legal structure of a disability insurance scheme should be designed such that it minimizes screening error. This calls for unambiguous definition of the disability risk and transparent assessment procedures to guarantee equal treatment of claimants. Moreover, eligibility standards should be strict, and should focus on residual capacity, not on *incapacity*. Permanent benefits can only be obtained after thorough examination of rehabilitative potential.

The expertise to do this can be obtained from those involved in work injury programs. Except the Netherlands, all countries distinguish between work-related and general disabilities. Work injury programs do not only show the advantages of strict definition and stressing rehabilitation but also those of private market supply of social insurance. Hence, the second set of instruments derives from the efficiency rule that risks should be pooled at the level where they can be best influenced. The lower the level of risk sharing the larger the returns to avoiding or limiting damages.

Under mandatory insurance the lowest possible pooling base is the firm. Mandating firms to carry the work-related disability risk of their personnel is quite common. But it is also defensible to charge the employer with carrying the general disability risks. It creates an

incentive structure in which employers efficiently weigh investments in prevention and job accommodation against lower benefit expenses, or insurance premiums.

Similar incentive structures can be found in countries like Argentina, Chile, and Switzerland that make their general disability insurance scheme part of their multi-pillar pension arrangement. Pooling disability risks in a pension fund may provide weaker incentives to use disability insurance carefully than charging the employer. But it has also major advantages. First, it combines redistribution (income solidarity), and inflation proof benefits, in the first pillar with individual responsibility, and demographic neutrality, in the second pillar.

Second, by making disability insurance part of the pension system the emphasis is on rehabilitation and permanent pensions are only awarded to those without prospects of recovery.

Third, having disability and old-age pensions financed and administered by the same funds (AFPs) allows for efficient insurance of permanent disability benefits. Permanent disability benefits are funded by a pension (savings) account. If the accumulated capital is too small to pay out the statutory disability benefit the insurance company adds a lump sum. The fact that disability risks increase with age but the additional sum needed to fund disability benefits up to pension age decreases with age makes for a limited and calculable risk. Mature private insurance markets have no problems covering such risks. This allows profiting from the efficiency gains engendered by delivery of social insurance through competition, and from economies of scope induced by combining two negatively correlated risks: disability and longevity.

Other countries, such as Latvia and Poland, have recently adopted multi-pillar pension systems. Their disability insurance is not part of this system. Ignoring political constraints we would recommend integration of the two schemes into the multi-pillar format.

## References

- Aarts, Leo, and Philip de Jong. 1996. *Private voorziening van sociale zekerheid in de praktijk. Een empirische studie naar de verdeling van private en publieke verantwoordelijkheden in de verzekering van het risico van beroepsgebonden arbeidsongeschiktheid in de Verenigde Staten*, COSZ/Ministerie van Sociale Zaken en Werkgelegenheid, The Hague: VUGA.
- Aarts, Leo, and Philip de Jong. 1998. "Privatization of Social Insurance and Welfare State Efficiency: Evidence from the Netherlands and the United States", in: Peter Flora et al., *The State of Social Welfare 1997*, Aldershot U.K.: Ashgate.
- Aarts, Leo, Richard Burkhauser, and Philip de Jong. 1996. *Curing the Dutch Disease. An International Perspective on Disability Policy Reform*, Aldershot U.K.: Avebury.
- Aarts, Leo, Richard Burkhauser, and Philip de Jong. 1998. "Convergence: A Comparison of European and United States Disability Policy", in: Terry Thomason, John F. Burton, Jr., and Douglas E. Hyatt, *New Approaches to Disability in the Workplace*, Industrial Relations Association Series: 299 - 388.
- Barr, Nicholas. 1998. *The Economics of Welfare State*, 3d Edition, London: Weidenfeld & Nicholson.
- Butler, Richard. 1994. "Safety Incentives in Workers' Compensation", in John F. Burton, Timothy P. Schmidle, *1995 Workers' Compensation Year Book*, Horsham (PA): I-82-91.
- Diamond, Peter, and Eytan Sheshinski. 1995. "Economic Aspects of Optimal Disability Benefits", *Journal of Public Economics*: 1-23.
- Fischback, P.V., and S.E. Kantor. 1995. "Did Workers Pay for the Passage of Workers' Compensation Laws?", *Quarterly Journal of Economics*: 713-42.
- Howards, Irving, Henry P. Brehm, and Saad Z. Nagi. 1980. *Disability from Social Problem to Federal Program*, New York: Praeger
- Lindbeck, Assar. 1994. "Uncertainty under the Welfare State: Policy-induced Risk", OCFEB Research Memorandum 9403, Rotterdam: Erasmus University Rotterdam.
- Nagi, Saad Z. 1969. *Disability and Rehabilitation: Legal, Clinical, and Self-Concepts and Measurements*, Ohio: Ohio State University Press.
- Parsons, Donald O. 1991. "Self-Screening in Targeted Public transfer Programs", *Journal of Political Economy*: 859-876.
- Rupp, Kalman, and David Stapleton. 1995. "Determinants of the Growth in Social Security Administration's Disability Programs – an Overview", *Social Security Bulletin*: 43-70.
- Shavell, Steven. 1979. "On Moral Hazard and Insurance", *Quarterly Journal of Economics*: 541-562.
- World Bank. 1994. *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth*. A World Bank Policy Research Report. Oxford university Press.
- World Health Organization. 1980. *International Classification of Impairments, Disabilities and Handicaps*, Geneva: WHO.

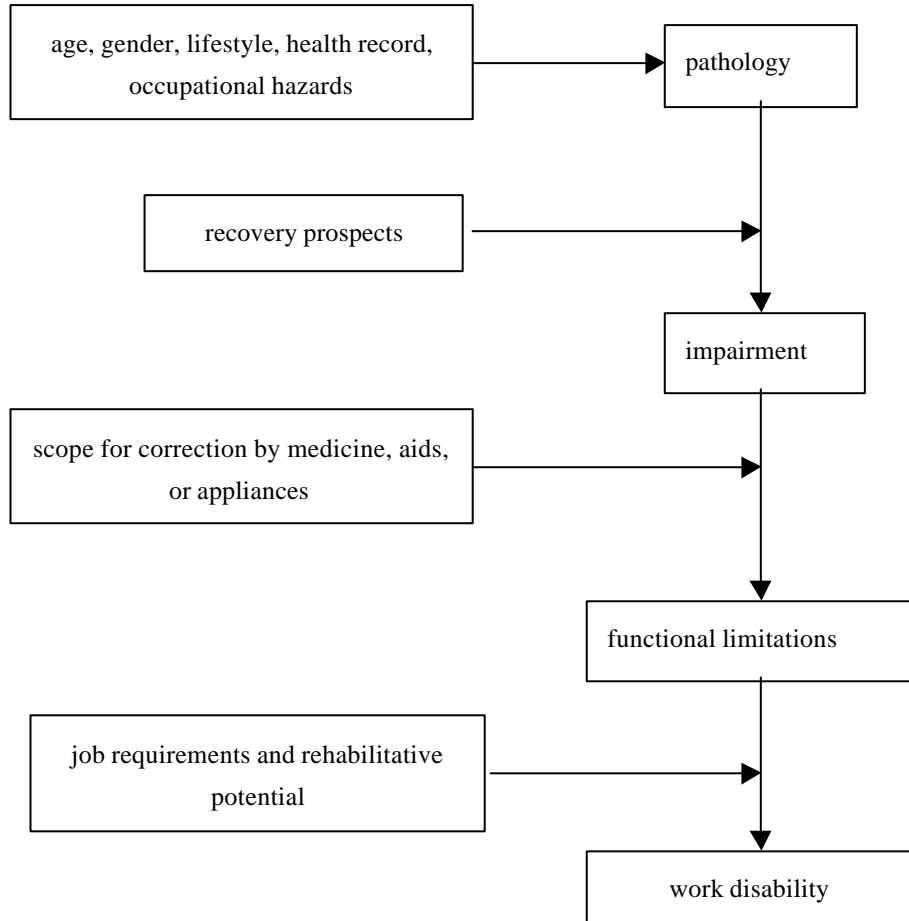
**Table 1**      **Disability Transfer Recipients Per Thousand Labor Force Participants**  
**(1995)**

United States <sup>(1)</sup>	64
Argentina <sup>(2)</sup>	
Chile <sup>(3)</sup>	7
United Kingdom <sup>(4)</sup>	68
Netherlands	142
Sweden	106
Norway	101
Germany, total	65
West	47
East	113
Austria	66
Switzerland	44
Poland	200
Latvia	100
Hungary	100

- (1) Included are Social Security Disability Insurance, and Supplemental Security Income; benefit recipients.
- (2) No data are available about beneficiaries from the old (paygo) system.
- (3) Beneficiaries per 1,000 current contributors to the private pension system.
- (4) Data are from 1991.

*Scheme 1 Etiology of Work Disability*

---



*Scheme 2 Social versus Private Insurance*

	<b>Social</b>	<b>Private</b>
Participation	Mandatory	Voluntary
Policy conditions	Uniform	Negotiable
Premium rates	Solidaristic	Risk-dependent
Financing method	Pay-as-you-go	Funding
Insurance provider	Public monopoly	Private competitors
Claim management	Public monopoly	Private competitors