Disability Risk, Disability Insurance and Life Cycle Behavior

Hamish Low
University of Cambridge and IFS

Luigi Pistaferri
Stanford University

October 29, 2008

Abstract

We provide a framework for understanding how disability risk and features of the disability insurance program in the US affect individual consumption and labor supply behavior in an explicit life-cycle setting. We decompose earnings risk into disability shocks (that reduce the ability to work) and shocks to general productivity. We identify structural parameters of the model using indirect inference and longitudinal data on consumption, disability status, disability insurance receipt, and earnings. We use our model to evaluate both the insurance benefit and the incentive effects of various program changes: (a) Increasing the "strictness" of the disability test; (b) Changing the probability of re-assessment for DI; (c) Changing the progressivity of DI payments, and (d) Reducing means-tested benefits that provide a consumption floor to DI applicants. We show that increasing the strictness of screening would increase welfare.

PRELIMINARY AND INCOMPLETE

1 Introduction

In the last 20 years the proportion of Disability Insurance (DI) claimants in the US has almost doubled (from about 2.5% to almost 5% of the population) and the proportion of Social Security spending taken up by the DI program has risen from 9.9% to 16.7% (see Figure 1 and the discussion in Autor and Duggan, 2006). These trends have been cited as an explanation for the decline in labor market participation of men and they have important implications for the long-term sustainability of the Social Security system. There is an underlying concern that the DI program is now being used as a gateway for early retirement, thus contradicting the original purpose of the program of providing insurance against rare but serious adverse health shocks. To evaluate these concerns and to evaluate the costs and benefits of changing the DI program to try to reduce disincentives to work, we need a realistic framework that models both the insurance benefit of DI as well as the incentive effects on individual choices over the life-cycle about labour supply, saving and application for DI. The underlying aim of this paper is to provide this quantitative evaluation of the DI program in an explicit life-cycle setting.

This paper has three specific goals. First, we propose a theoretical framework that allows us to study the effect of disability risk on behavior in an integrated framework, i.e., modeling life-cycle labor supply, savings and the DI application decisions jointly. This framework is both general and realistic. We consider the problem of an individual who faces two types of shock to wages. The first is a permanent productivity shock unrelated to health. The second is a "disability" shock which reduces the ability to work. The distinction between the two types of shock to wages is key

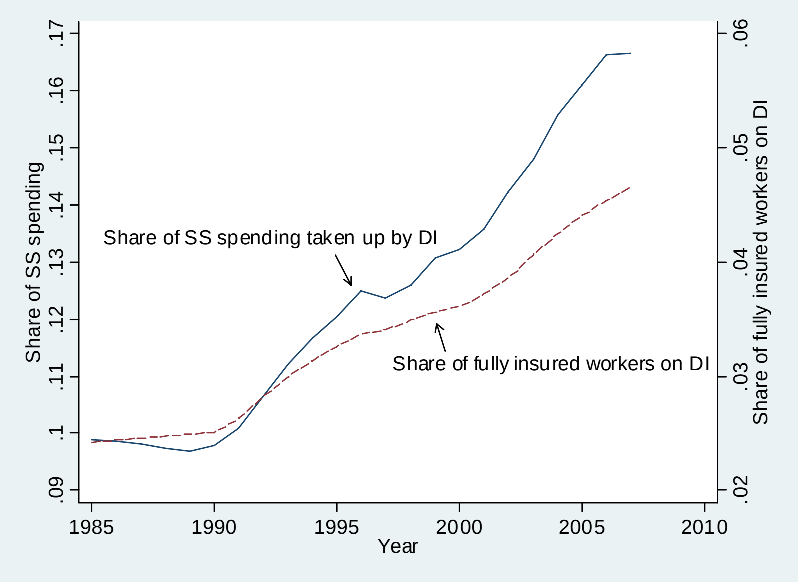

Figure 1: The growth of the DI program, 1985-2007

for understanding the *moral hazard* problem with the DI program.[1] Individuals with a disability shock above a certain threshold can not work. Individuals with a productivity shock below a certain threshold may not want to work and apply for DI benefits. Whether such labor supply distortions occur depends on a number of factors, such as the extent of labor market frictions and the availability of alternative forms of insurance (own savings, as well as other government-provided insurance programs).

Second, we estimate the relevant structural parameters within the context of our model. We use PSID data on wages and indicators of disability status to help identify the parameters of the wage process and data on consumption loss to identify preference parameters. Finally, we identify the structural policy parameters governing the disability application and review process using data on the stock of individuals on DI, the flows onto DI, and the flows off DI separately by disability status.

Third, we want to use our model and the estimates of the structural parameters to tackle the welfare and policy questions more directly. We address a number of questions: first, how well insured are individuals against disability risk; second, how responsive are labour supply and savings to changes in the details of the DI program; third, does the loss of insurance associated with tightening the criterion for DI offset the benefit in terms of reduced false applications; fourth, would an asset test on DI have a beneficial effect on expected utility; finally, are there important interactions between different government programs: for example, the presence of food stamps provides a floor to consumption and the level of this floor may affect applications for DI. The ability to evaluate these questions in a coherent unified framework is one of the main benefits of the paper.

Some of these issues have been addressed elsewhere in the literature. The implications of DI

[1] We use the term "moral hazard", even though there is no hidden action, to be consistent with the terminology adopted in, among others, Bound and Burkhauser (1999) and popular public finance textbooks, such as Gruber (2005).

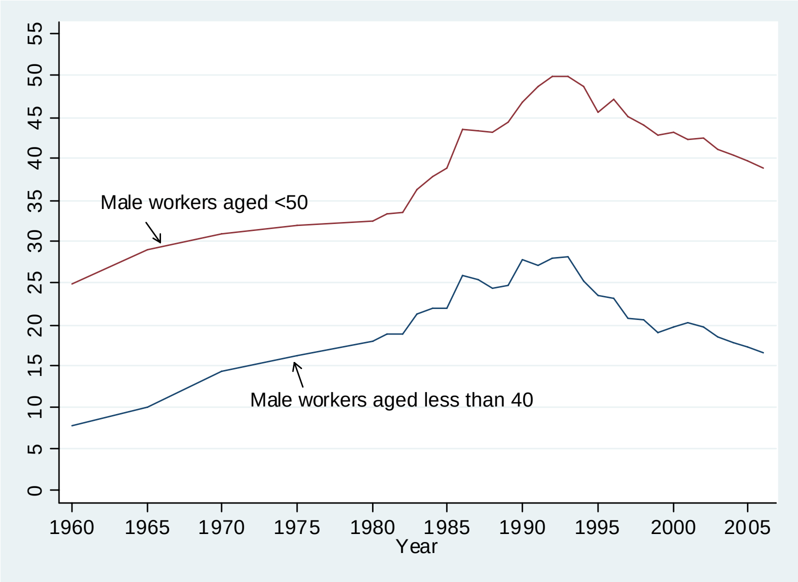

Figure 2: Proportion of new DI awards

for labour supply was first studied by Parsons (1980) who showed the correlation between the rise in applications for DI and the fall in male labour force participation. The apparent disincentive effects of DI were supported by Gruber (2000) and Bound et al. (2001). Kreider (1999) uses a more structural approach to understand the joint decision of applying for DI and of not participating in the labour market, and finds that although DI has important disincentive effects on labour supply, the change in DI generosity cannot explain the fall in labour force participation. Autor and Duggan (2007) provide a detailed analysis of the trends in DI receipt and the causes behind these trends and conclude that the growth is due to more generous benefits and more lax screening. This is consistent with the structural analysis of Benitez-Silva et al. (2007) who use a structural model to assess the effectiveness of the DI screening process in the 1990s. Their conclusion is that over 40% of recipients of DI are not truly work limited and this adds to the picture of an inefficient insurance program.

Most of the structural analyses of DI errors have used HRS data. The HRS has the advantage over the PSID of asking very detailed questions on disability status and insurance, minimizing measurement error. However, the HRS samples from a population of older workers and retirees (aged 50 or more). This is an important limitation. As shown in Figure 2, young male workers (defined as those younger than 40) account for one-fifth to one-fourth of the flows of new entrants in the Disability Insurance program in recent years. If we extend the definition of "young worker" to include those younger than 50, we find that between 40% and 50% of new entrants in the program are "young". Data sets like the HRS, which purposely samples cohorts older than 50, would miss completely this phenomenon. From a policy perspective, it is very important to understand what brings young workers to apply for and receive Disability Insurance benefits. In fact, young workers who are on DI contribute less to its financing and are expected to stay longer on the program (at least for disabilities not leading to death). One of the advantages of PSID is that it allows us to study the entire life-cycle rather than parts of it.

The broader issue of the value of DI requires an evaluation of the benefit of the insurance provided by DI as well as an assessment of the efficiency loss. Bound et al. (2004) carry out some calculations of welfare costs of disability insurance, but without using a full model of behaviour. Such a framework is however necessary to evaluate the effectiveness and net benefit of proposed reforms.

One such proposal is the Golosov and Tsyvinski (2006) proposal to impose an asset test on disability applicants. Golosov and Tsyvinski show that, if disability status is private information, an asset test can implement the constrained Pareto optimum. This result may depend on the assumption that assets are used to smooth periods of non-employment associated with disability. In our framework, where assets held for precautionary reasons are substitutable with assets held for life-cycle reasons, the welfare benefit of an asset test is ambiguous. Alternative proposals include increasing the medical hurdle for applicants, raising the reassessment rate among recipients, increasing the waiting time before a DI application is permitted. The analysis of policy in a life-cycle framework as in Hubbard et al. (1995) has not, however, explicitly considered health risks, which may differ in important ways from productivity risk, or modelled the disability insurance program.

There have been some recent papers identifying the extent of health risk. In particular, DeNardi, French and Jones (2006) estimate the risk to health expenditure, but their focus is on the elderly, rather than those of working age when disability insurance is active. Adda, Banks and Gaudecker (2006) estimate the effect of income shocks on health and find only small effects. Meyer and Mok (2007) and Stephens (2001) estimate in a reduced form way the effect of earnings losses associated with disability on consumption. The value of our paper is in combining estimates of the risk associated with health shocks in a framework that allows the evaluation of the social insurance provided by DI.

Section 2 presents the life-cycle model allowing for health status, and discusses the various social insurance programs available to individuals. Section 3 summarises the data used in the estimation of the model, focusing on the data on disability status and on consumption. Section 4 discusses the identification strategy. Section 5 presents the estimates of the structural parameters. Section 6 discusses the implications of the results for the optimality of the parameters of the disability insurance program and section 7 concludes.

2 Life-Cycle Model

2.1 Individual Problem

We consider an individual with a period utility function

$$U_t = U(c_t, P_t; L_t)$$

where P_t is a discrete $\{0,1\}$ labor supply participation variable, c_t consumption and L_t is a discrete disability status indicator $\{0,1,2\}$. The individual is assumed to maximize lifetime expected utility

$$\max_{c,P,DI^{App}} V_t = E_t \sum_{s=t}^{T} \beta^{s-t} U(c_s, P_s; L_s)$$

where β is the discount factor and E_t the expectations operator conditional on information available in period t (a period being a quarter of a year). Individuals live for T periods, may work from age

22 to 62, and face an exogenous mandatory spell of retirement of 10 years at the end of life. The date of death is known with certainty.

The intertemporal budget constraint during the working life has the form

$$A_{t+1} = R \begin{bmatrix} A_t + \left(w_t h \left(1-\tau_w\right) - F\left(L_t\right)\right) P_t \\ +(B_t E_t^{UI}\left(1-E_t^{DI}\right) + DI_t E_t^{DI})\left(1-P_t\right) \\ +W_t E_t^W - c_t \end{bmatrix}$$

where A are beginning of period assets, R is the interest factor, w the hourly wage rate, h a fixed number of hours (corresponding to 500 hours per quarter), τ_w a proportional tax rate that is used to finance social insurance programs, F the fixed cost of work that depends on disability status, B_t unemployment benefits, W_t the monetary value of the means tested welfare payment, DI_t the amount of disability insurance payments obtained, and E_t^{UI}, E_t^{DI}, and E_t^W are recipiency $\{0,1\}$ indicators for unemployment insurance, disability insurance, and the means-tested welfare program, respectively.

The worker's problem is to decide whether to work or not. When unemployed he has to decide whether to accept a job that may have been offered or wait longer. If eligible, the unemployed person will have the option to apply for disability insurance. Whether employed or not, the individual has to decide how much to save and consume. Accumulated savings can be used to finance spells out of work and retirement.

We use a utility function of the form

$$u\left(c_t, P_t; L_t\right) = \frac{\left(c_t \exp\left(\theta L_t\right) \exp\left(\eta P_t\right)\right)^{1-\gamma}}{1-\gamma}$$

We impose that $\gamma > 1$.The parameter θ captures the utility loss for the disabled in terms of consumption ($\theta < 0$). Participation also induces a utility loss determined by the value of η ($\eta < 0$). This implies that consumption and participation are Frisch complements (i.e. the marginal utility of consumption is higher when participating) and that the marginal utility of consumption is higher when suffering from a work limitation.

We assume that individuals are unable to borrow:

$$A_t \geq 0$$

In practice, this constraint has bite because it precludes borrowing against unemployment insurance, against disability insurance, against social security and against the means-tested program.

At retirement, people collect social security benefits which are paid according to a formula similar to the one we observe in reality (see below). These benefits, along with assets that people have voluntarily accumulated over their working years, are used to finance consumption during retirement.

2.2 The Wage Process and Labour Market Frictions

We model the wage process for individual i as being subject to general productivity shocks and shocks to the disability status (as well as the contribution of observable characteristics X_{it}):

$$\ln w_{it} = X_{it}'\alpha + \beta_1 \mathbf{1}\left\{L_{it} = 1\right\} + \beta_2 \mathbf{1}\left\{L_{it} = 2\right\} + \varepsilon_{it} + \omega_{it} \tag{1}$$

where ω_{it} is an i.i.d. measurement error, and

$$\varepsilon_{it} = \varepsilon_{it-1} + \eta_{it}$$

We make the assumption that the two shocks η_{it} and ω_{it} are independent and that disability shocks are orthogonal to the general stochastic component of individual productivity. Our goal is to identify the variance of the productivity shock σ_{η}^{2} as well as β_1 and β_2.

Individuals work limitation status, L_{it}, evolves according to a three state first-order Markov process which is age dependent. Upon entry into the labor market, all individuals are assumed to be healthy ($L_{i0} = 0$). Transition probabilities from any state depend on age. We assume that these transition probabilities are exogenous and in particular, we rule out the possibility of individuals investing in health prevention.[2]

Equation (1) determines the evolution of individual productivity. Productivity determines the offered wage when individuals receive a job offer. In our framework, individuals make a choice about whether or not to accept an offered wage. This will also depend on the fixed costs of work, which in turn depend on the extent of the work limitation, $F(L)$. In addition, there are labour market frictions which mean that not all individuals receive job offers. First, there is job destruction, δ, which forces individuals into unemployment for (at least) one period. Second, job offers for the unemployed arrive at a rate λ and so individuals may remain unemployed.

This wage and employment environment implies a number of sources of risk, from individual productivity, work limitation shocks and from market frictions. These risks are idiosyncratic, but we assume that there are no markets to provide insurance against these risks. Instead, there is partial insurance coming from government insurance programs (as detailed in the next section) and from individuals' own saving.

2.3 Social Insurance

2.3.1 The SSDI Program

The Social Security Disability Insurance program (SSDI) is an insurance program for covered workers, their spouses, and dependents that pays benefits related to average earnings. The purpose of the program is to provide insurance against health shocks that impair substantially the ability to work. The difficulty with providing this insurance is that health status and the impact of health on the ability to work is imperfectly observed.[3]

The program was enacted in 1956 for individuals older than 50 and suffering from an impairment that was "expected to result in death or be of long, continued, and indefinite duration". In later years eligibility was extended to individuals under age 50, disability did not have to be permanent any more, waiting periods were reduced and benefit levels increased. By the mid-1970s typical after-tax replacement rates reached 60%. The Social Security Administration (SSA) responded to an enormous growth in the DI roll by refining their regulations guiding decisions and by changing the frequency and nature of medical eligibility reviews for DI beneficiaries, which lead to a fall in award rates from 48.8% to 33.3% between 1975 and 1980 and to an increase in the number of terminations. In the 1984, eligibility criteria were liberalized, when the SSA issued new rulings that gave controlling weight to source evidence (e.g., own physician).

The award of disability insurance now depends on the following conditions:

[2] We allow the process to differ by education, which may implicitly capture differences in health investments.

[3] About 25% of workers in the private sector are covered by employer-sponsored long-term disability insurance plans.

1. An individual has to have filed an application for disabled worker's benefits.
2. There is a work requirement on the number of quarters of prior participation: Workers over the age of 31 are disability-insured if they have 20 quarters of coverage during the previous 40 quarters.
3. There is a statutory five-month waiting period out of the labour force before an application will be processed.
4. Finally, the individual must meet a medical requirement, i.e. the presence of a disability defined as the:

   > Inability to engage in any substantial gainful activity by reason of any medically determinable physical or mental impairment, which can be expected to result in death, or which has lasted, or can be expected to last, for a continuous period of at least 12 months

This requires that the disability affects the ability to work; and further, both the severity and the expected persistence of the disability matter. Figure 3 gives a graphical representation of the sequential DI determination process. After determining the presence of a medical disability (step 2) that is not a listed impairment (step 3),[4] the DI evaluators try to determine whether the disability is a *working* disability. The last step in the sequence connects the "pathological" criterion with an "economic" criterion. Two individuals with the identical working disability may receive different DI determination decisions depending on the economic conditions they face at the time the determination is made.

In our model, we make the following assumptions in order to capture the main aspects of the disability insurance program detailed above:

1. Individuals have to make the choice to apply for benefits.
2. Individuals have to have been at work for at least one period prior to becoming unemployed and making the application.
3. Individuals must have been unemployed for at least one quarter before applying. Successful applicants begin receiving benefits in that second quarter. Unsuccessful individuals must wait a further quarter before being able to return to work, but there is no direct monetary cost of appplying for DI.
4. The probability of success depends on the true work limitation status and on age.

$$\Pr\left(DI_t = 1 |\, DI_t^{App} = 1, L_t, t\right) = \begin{cases} P_L^{S,Y} & if\ t < 45 \\ P_L^{S,O} & if\ 62 \geq t \geq 45 \end{cases}$$

The medical requirement in the SSDI program imposes a severity and persistence requirement on the work limitation. In our model, the expected persistence of the work limitation is captured by the Markov process for wages and is age dependent. This age dependence of the

[4] The listed impairments are described in "Disability Evaluation under Social Security", a blue-book published periodically by the SSA. The listed impairments are physical and mental conditions for which specific disability approval criteria has been set forth or listed (for example, chronic asthmatic bronchitis or amputation of limbs).

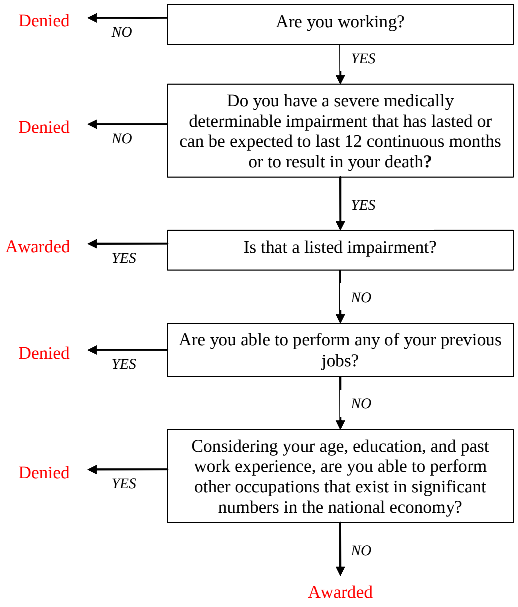
Are you working?
NO
Denied
YES
Do you have a severe medically determinable impairment that has lasted or can be expected to last 12 continuous months or to result in your death?
NO
Denied
YES
Is that a listed impairment?
YES
Awarded
NO
Are you able to perform any of your previous jobs?
YES
Denied
NO
Considering your age, education, and past work experience, are you able to perform other occupations that exist in significant numbers in the national economy?
YES
Denied
NO
Awarded

Figure 3: Disability Insurance Determination

persistence in our model is the reason why we make the probability of a successful application for DI dependent on age. The survey question survey we use (described below) matches this criterion because it asks individuals about work limitations. Finally, we account for the "economic" role played by the last step of the DI determination process by adding labor market frictions to our model.

Individuals leave the disability program either either voluntarily (which in practice means into employment) or following a reassessment of the work limitation and being found to be able to work. The probability of being reassesed is 0 for the first year, then is given by P^{Re}, which is independent of L and age. If an individual is not successful on application or if an individual is rejected on reassessment, the individual has to remain unemployed until the next quarter before taking up a job. Individuals can only re-apply in a subsequent unemployment spell.

SSDI benefits are calculated in essentially the same fashion as Social Security retirement benefits, and have been subject to the same changes in benefit levels. Beneficiaries receive indexed monthly payments equal to their Primary Insurance Amount (PIA), which is based on taxable earnings averaged over the number of years worked. Caps on the amount that individuals can receive from (or pay into) the DI system make the system progressive. Because of the progressivity of the benefits and because of the fact that individuals receiving SSDI also receive Medicare benefits after two years, the replacement rates (i.e. the percentage of before-disability income an individual will receive once she ceases working) are substantially higher for workers with low earnings and those without employer-provided health insurance. However, benefits are independent of the extent of the work limitation.

In the model, we set the value of the benefits according to the actual schedule in the US program. The value of disability insurance is given by

$$D_{it} = \begin{cases} 0.9 \times \overline{w}_i & \text{if } \overline{w}_i \leq a_1 \\ 0.9 \times a_1 + 0.32 \times (\overline{w}_i - a_1) & \text{if } a_1 < \overline{w}_i \leq a_2 \\ 0.9 \times a_1 + 0.32 \times (a_2 - a_1) + 0.15 \times (\overline{w}_i - a_2) & \text{if } a_2 < \overline{w}_i \leq a_3 \\ 0.9 \times a_1 + 0.32 \times (a_2 - a_1) + 0.15\,(a_3 - a_2) & \text{if } \overline{w}_i > a_3 \end{cases} \tag{2}$$

where $\overline{w}_i$ is average earnings computed before the time of the application and a_1, a_2, and a_3 are thresholds we take from the legislation.[5] We assume $\overline{w}_i$ can be approximated by the value of the permanent wage at the time of the application. Whether an individual is eligible (i.e., $E_{it}^{DI} = 1$) depends on the decision to apply ($DI_{it} = 1$) while being out of work and on having received a large negative productivity shock. We assume that the probability of success is independent of age. Eligibility does not depend on whether an individual quits or the job is destroyed.

In retirement, all individuals receive social security calculated using the same formula used for disability insurance.

2.3.2 Unemployment Insurance

We assume that unemployment benefits are paid only for the quarter immediately following job destruction. We define eligibility for unemployment insurance E_{it}^{UI} to mirror current legislation: benefits are paid only to people who have worked in the previous period, and only to those who had their job destroyed (job quitters are therefore ineligible for UI payments, and we assume this can be perfectly monitored).[6] We assume $B_{it} = b \times w_{it-1}\overline{h}$, subject to a cap, and we set the replacement ratio $b = 75\%$. This replacement ratio is set at this high value because the payment that is made is intended to be of a similar magnitude to the maximum available to someone becoming unemployed.

In the US, unemployment benefit provides insurance against job loss and insurance against not finding a new job. However, under current legislation benefits are only provided up to 26 weeks (corresponding to two periods of our model) and so insurance against not finding a new job is limited. Our assumption is that there is no insurance against the possibility of not receiving a job offer after job loss. This simplifying assumption means that, since the period of choice is one quarter, unemployment benefit is like a lump-sum payment to those who exogenously lose their job and so does not distort the choice about whether or not to accept a new job offer. The only distortion is introduced by the tax on wages.

2.3.3 Universal Means-Tested Program

In modelling the universal means-tested program, our intention was to mirror partially the actual food stamps program but with three important differences. First, the means-testing is only on household income rather than on income and assets; second, the program provides a cash benefit rather than a benefit in kind; and third, we assume there is 100% take-up.[7] These assumptions

[5] In reality what is capped is $\overline{w}_i$ (PIA). We translate a cap on PIA into a cap on DI payments.

[6] We have simplified considerably the actual eligibility rules observed in the US. A majority of states have eligibility rules which are tougher than the rule we impose, both in terms of the number of quarters necessary to be eligible for any UI and in terms of the number of quarters of work necessary to be eligible for the maximum duration (Meyer, 2002). However, making eligibility more stringent in our model is numerically difficult because the history of employment would become a state variable. Our assumption on eligibility shows UI in its most generous light.

[7] The difficulty with allowing for an asset test in our model is that there is only one sort of asset which individuals use for retirement saving as well as for short-term smoothing. In reality, the asset test applies only to liquid wealth and thus excludes pension wealth (as well as real estate wealth and other durables). We are working on a new version of the model that relaxes this. Moreover, we introduce an exogenous demographic life-cycle which changes the value

mean the program plays the role of providing a floor to income for all individuals. This is similar to Hubbard, Skinner and Zeldes (1995). Gross income is given by

$$y_{it}^{gross} = w_{it}hP_{it} + \left(B_{it}E_{it}^{UI}\left(1 - E_{it}^{DI}\right) + D_{it}E_{it}^{DI}\right)(1 - P_{it}) \tag{3}$$

giving net income as $y = (1 - \tau_w)\, y^{gross} - d$, where d is the standard deduction that people are entitled to when computing net income for the purpose of determining food stamp allowances. The value of the program is then given by

$$T_{it} = \begin{cases} \overline{T} - 0.3 \times y_{it} & \text{if } E_{it}^{T} = 1 \ \left(\text{i.e., if } y_{it} \leq \underline{y}\right) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The maximum value of the payment, $\overline{T}$, is set assuming a household with two adults and two children, although in our model there is only one earner. The term $\underline{y}$ should be interpreted as a poverty line. In the actual food stamp program, only people with net earnings below the poverty line are eligible for benefits ($E_{it}^{T} = 1$).

The distinction with the actual food stamps program is that the means-tested program in this paper is not asset tested. The program interacts in complex ways with disability insurance: the Food Stamps program provides a consumption floor during application for DI.

2.3.4 Taxation on Earnings

The tax rate on earnings, τ_w, is set to hold the government budget in balance when varying the parameters of the social insurance policies.

2.4 Model Discussion

Our characterisation of the application process and the trade-off between genuine applicants and non-genuine applicants is represented qualitatively in figure 4 (assuming work limitation is a continuous variable for simplicity). There is a threshold level of L which is intended as the cut-off for eligibility for DI. This stringency threshold, $\overline{S}$, is the level of disability at which the government judges that no gainful employment is possible, and is independent of the productivity level. Below this level of L, some individuals may wish to apply for DI if their productivity is sufficiently low because the government only observes a noisy measure of the true disability status. Figure 4 shows the threshold value of productivity which determines whether an individual chooses to try to "cheat" the system and apply for DI. A lower level of disability means income is not affected significantly by disability and that there is less chance of being wrongly assessed as needing DI and this implies that applications for DI will be made only by individuals with a very low level of productivity. Further, the opportunity cost of applying is greater if income is higher and those in better health have higher incomes. This decision to apply will depend on assets, age and other characteristics.

Benitez-Silva et al. (2006) characterise in a very compelling way the extent of moral hazard in disability insurance applications. In particular, they show that 40% of recipients do not conform to the criterion of the SSA. This raises the question of whether the “cheaters” are not at all disabled or whether they have only a partial disability. With our characterisation of individuals as falling

of Food Stamps benefits depending on the head’s age. Finally, we introduce a program that mirrors SSI. We assume that people who are on DI and have income below a certain (poverty) threshold, also receive SSI from the government.

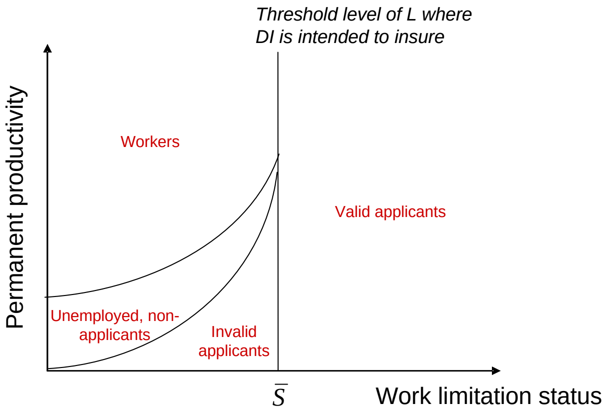

Figure 4: Valid and Invalid Applications for DI

into categories severely restricted ($L = 2$) and at least partially restricted ($L = 1$), we are able to explore this issue.

The criteria quoted above specifies "any substantial gainful activity": this refers to a labour supply issue. However, it does not address the labour demand problem. Of course, if the labour market is competitive this will not be an issue because workers can be paid their marginal product whatever their productivity level. In the presence of imperfections, however, the wage rate associated with a job may be above the disabled individual's marginal productivity. The Americans with Disability Act (1992) tries to address this question but that tackles the issue only for incumbents who become disabled.

2.5 Solution

There is no analytical solution for our model. Instead, the model must be solved numerically, beginning with the terminal condition on assets, and iterating backwards, solving at each age for the value functions conditional on work status. The solution method is discussed in more detail in the appendix. Here we describe the main features of the algorithm used.

We start by constructing the value functions for the individual when employed and when out of work. When employed, the state variables are $\{A_{it}, \varepsilon_{it}, L_{it}\}$, corresponding to current assets, individual productivity and health status. We denote the value function when employed as V^e. When unemployed, there are three alternative discrete states the individual can be: unemployed and not applying for disability (giving a value V^n), unemployed and applying for disability (giving a value V^{App}), and unemployed and already receiving disability insurance (giving a value V^{Succ}). We consider the specification of each of these value functions in turn.

Value function if working:

$$V_t^e\left(A_{it},\varepsilon_{it},L_{it}\right) = \max_{c}\left\{\begin{array}{c} U\left(c_{it},P_{it}=1;L_{it}\right)+ \\ \\ \beta\delta E_t\left[V_{t+1}^n\left(A_{it+1},\varepsilon_{it+1},L_{it+1},DI_{it+1}^{Elig}=1\right)\right] \\ \\ +\beta\left(1-\delta\right)E_t\max\left\{\begin{array}{c} V_{t+1}^n\left(A_{it+1},\varepsilon_{it+1},L_{it+1},DI_{it+1}^{Elig}=1\right) \\ V_{t+1}^e\left(A_{it+1},\varepsilon_{it+1},L_{it+1}\right)\end{array}\right\}\end{array}\right.$$

We consider now the value function for an unemployed individual who is not applying for disability insurance in period t. We need to define as a state variable whether or not an individual has already applied for disability in the current unemployment spell in order to distinguish between those who have the option of applying for disability and those who are ineligible to apply. The value function when eligible for disability is given by:

$$V_t^n\left(A_{it},\varepsilon_{it},L_{it},DI_t^{Elig}=1\right) = \max_{c,DI^{App}}\left\{\begin{array}{c} u\left(c_{it},P_{it}=0;L_{it}\right) \\ \\ +\beta\left\{1.App=1\right\}E_tV_{t+1}^{App}\left(A_{it+1},\varepsilon_{it+1},L_{it+1}\right) \\ \\ +\beta\left\{1.App=0\right\} \\ \left[\begin{array}{c} \lambda^nE_t\max\left\{\begin{array}{c} V_{t+1}^n\left(A_{it+1},\varepsilon_{it+1},L_{it+1},DI_{t+1}^{Elig}=1\right) \\ V_{t+1}^e\left(A_{it+1},\varepsilon_{it+1},L_{it+1}\right)\end{array}\right\} \\ +\left(1-\lambda^n\right)E_t\left[V_{t+1}^n\left(A_{it+1},\varepsilon_{it+1},L_{it+1},DI_{t+1}^{Elig}=1\right)\right]\end{array}\right]\end{array}\right.$$

The value function when applying is given by

$$V_t^{App}\left(A_{it},\varepsilon_{it},L_{it}\right) = \max_{c}\left\{\begin{array}{c} u\left(c_{it},P_{it}=0;L_{it}\right) \\ \\ +\beta\Pr E_tV_{t+1}^{Succ}\left(A_{it+1},\varepsilon_{it+1},L_{it+1},D_t=0\right) \\ \\ +\beta(1-\Pr)E_tV_{t+1}^n\left(A_{it+1},\varepsilon_{it+1},L_{it+1},DI_{t+1}^{Elig}=0\right)\end{array}\right.$$

where D_t is a state variable for the duration of the spell on disability insurance and

$$\Pr = \Pr\left(\left.DI_t=1\right|DI_t^{App}=1,L_t\right).$$

Finally, we have to define the value function if an application for disability has been successful.

$$V_t^{Succ}\left(A_{it},\varepsilon_{it},L_{it},D_t\right)= \tag{5}$$

$$\max_c\left\{\begin{array}{c} u\left(c_{it},L_{it},P_{it}=0\right) \\ +\beta\left(1-\Pr\right)E_t\left[V_{t+1}^{Succ}\left(A_{it+1},\varepsilon_{it+1},L_{it+1},D_t+1\right)\right] \\ \beta\Pr E_t\left[V_{t+1}^{n}\left(A_{it+1},\varepsilon_{it+1},L_{it+1},DI^{Elig}=0\right)\right] \end{array}\right\} \tag{6}$$

where

$$\Pr=\Pr\left(\text{Reassessment}\right)\Pr\left(DI_t=0\left|L_t\right.\right)$$

Our model has discrete state variables for: Wage productivity, Work limitation status, Participation, Eligibility to apply for DI (if not working), and Length of time on DI (over 1 year or less than 1 year). The only continuous state variable is assets. We use backward induction to obtain policy functions.

Value functions are increasing in assets A_t but they are not necessarily concave, even if we condition on labor market status in t. The non-concavity arises because of changes in labor market status in future periods: the slope of the value function is given by the marginal utility of consumption, but this is not monotonic in the asset stock because consumption can decline as assets increase and expected labor market status in future periods changes. This problem is also discussed in Lentz and Tranaes (2001). By contrast, in Danforth (1979) employment is an absorbing state and so the conditional value function will be concave. Under certainty, the number of kinks in the conditional value function is given by the number of periods of life remaining. If there is enough uncertainty, then changes in work status in the future will be smoothed out leaving the expected value function concave: whether or not an individual will work in $t+1$ at a given A_t depends on the realization of shocks in $t+1$. Using uncertainty to avoid non-concavities is analogous to the use of lotteries elsewhere in the literature. In the value functions (**??**) and (**??**), the choice of participation status in $t+1$ is determined by the maximum of the conditional value functions in $t+1$.

2.6 Structural Parameters to Estimate

These are the structural parameters we want to estimate:

- Effect of disability on wages: β_1,β_2

  Disability risk (probability of having a work limitation in t, given past health)

  Productivity risk σ_η^2

  Labour market frictions: δ, $\lambda, F\left(L\right)$

- Probability of Success in DI application ($age<45$):

  $$P_{L=0}^{S,Y},P_{L=1}^{S,Y},P_{L=2}^{S,Y}$$

  Probability of Success in DI application ($age\geq45$):

  $$P_{L=0}^{S,O},P_{L=1}^{S,O},P_{L=2}^{S,O}$$

  Probability of Reassessment while on DI: P^{Re}

- Utility cost of a work limitation, θ

  Disutility of work, η

  Coefficient of relative risk aversion and the discount rate

3 Data

We conduct our empirical analysis using longitudinal data from the 1987-1993 Panel Study of Income Dynamics (PSID).[8] The PSID offers repeated, comparable annual data on disability status, disability insurance recipiency, earnings, and food consumption. Its main disadvantage is that the sample of people likely to have access to disability insurance is small and there may be some questions about the variables that define both disability status and disability insurance status (see below), especially in comparison to the definition of disability of the Social Security Administration.[9] The PSID sample we use excludes the Latino and SEO sub-sample, female heads, and people younger than 25 or older than 65. We are currently working with a much larger data set, extended to 2005, but have no results yet. The most important aspect of these new data is that starting in 1999 the PSID has added questions aimed at obtaining a more comprehensive measure of consumption (see Li et al, 2006).

3.1 Disability Data

We define a discrete indicator of work limitations (L_{it}), based on the following questions:

1. *Do you have any physical or nervous condition that limits the type of work or the amount of work you can do?*

To those answering "Yes", the interviewer then asks:

2. *Does this condition keep you from doing some types of work?*
   Possible answers are: "Yes", "No", or "Can do nothing".

To those who answer "Yes" or "No", the interviewer then asks:

3. *For work you can do, how much does it limit the amount of work you can do?*

Possible answers are: "A lot", "Somewhat", "Just a little", or "Not at all".

We distinguish between no work limitations ($L_{it} = 0$), moderate limitations ($L_{it} = 1$) and severe limitations ($L_{it} = 2$). We assume that an individual is affected by moderate work limitations if he answer "Yes" to the first question, "Yes" to the second and "Somewhat" to the third. We assume that an individual is affected by severe limitations if he answer "Yes" to the first question and "Can do nothing" to the second question, or if he answer "Yes" to the first question, "Yes" to the second and "A lot" to the third.

The validity of these self-reports is somewhat controversial for two reasons: first, individuals may over-estimate their work limitation in order to justify their disability payments or their non-participation in the labour force. Second, health status may be endogenous, and non-participation in the labour force may affect health (either positively or negatively). Regarding the first criticism,

[8] Due to the retrospective nature of the questions on earnings and consumption, this means our data refer to the 1986-1992 period.

[9] We considered using HRS instead of PSID. The HRS has the advantage over the PSID of asking very detailed questions on disability status and insurance, minimizing measurement error. However, the data on consumption in the HRS is asked to households who are too old (over 55) to benefit substantially from disability insurance (see Figure 2). Further, there is no strict alignment between the timing of the disability questions and the consumption questions. In particular the additional modules on consumption are conducted in 2001 and 2003 but the core questions are asked in 2000 and 2002, and these questions are asked to individuals born before 1947.

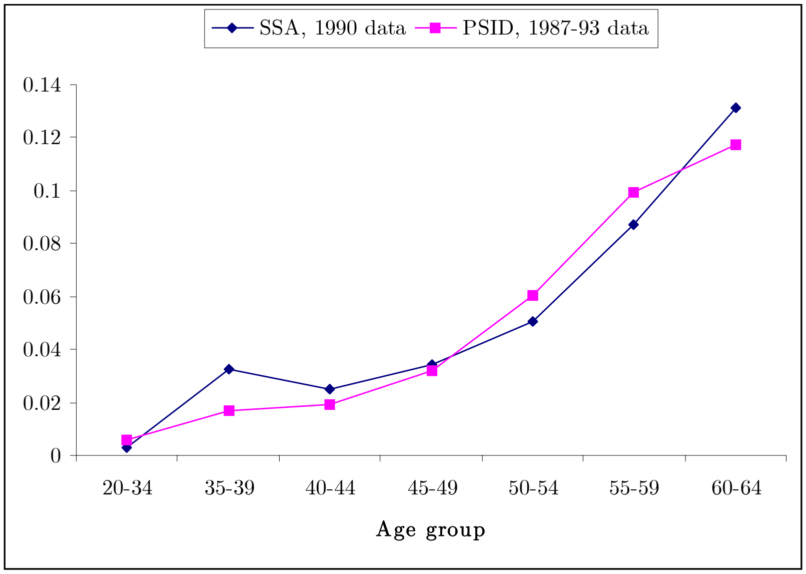

Figure 5:

Bound and Burkhauser (1999) survey a number of papers that show that self-reported measures are highly correlated with clinical measures of disability (Bound and Burkhauser, 1999). Benitez-Silva et al., (2003) show that self-reports are unbiased predictors of the definition of disability used by the SSA. Finally, Burkhauser and Daly (1996) show the validity of the PSID measures using the 1986 health supplement. Regarding the second criticism of the endogeneity of health status, Stern (1990) and Bound (1991) both find positive effects of non-pariticipation on health, but the effects are economically small. Further, Smith (2004) finds that income does not affect health once one controls for education.

3.2 Disability Insurance

To identify whether an individual in the PSID is receiving disability insurance, we use a question that asks whether the amount of social security payments received was due to disability.[10] This question is asked from 1986 onwards. Prior to 1986, the question was not targeted to the head of the household, and so we cannot distinguish the recipient of the insurance. Figure **??** shows how the fraction of individuals receiving DI increases with age. It also shows that the DI matches well proportions from the population. The match is good also in the time series. In the population, the proportion of people on DI has increased from 2.4% to 4.3% between 1985 and 2005. In the PSID the increase between 1986 and 2005 is from 2.4% to 4.5%.

[10] The survey first asks the amount of Social Security Payments Received in year t by the year $t+1$ head. Then, it asks "Was that disability, retirement, survivor's benefits, or what?". Possible responses are: 1) Disability, 2) Retirement, 3) Survivor's benefits; dependent of deceased recipient, 4) Dependent of disabled recipient, 5) Dependent of retired recipient, 6) Other, 7) Any combination of the codes above.

3.3 Consumption Data

One difficulty with the PSID is that the consumption in the data refers only to food. By contrast, in the model, the budget constraint imposes that over the lifetime, all income is spent on consumption. To compare consumption in the model to consumption in the data, we create non-durable consumption in the data by an imputation procedure. We estimate in the CEX:[11]

$$\ln c_{it} = \sum_{j=0}^{K} \theta_j \left(\ln F_{it}\right)^j + X_{it}' \mu + \varepsilon_{it}$$

We define

$$\widehat{\ln c}_{it} = \sum_{j=0}^{K} \widehat{\theta}_j \left(\ln F_{it}\right)^j + X_{it}' \widehat{\mu}$$

Using the imputed value for c_{it} increases the variance of our estimates of β_i but does not affect their consistency.[12]

4 Identification

Our identification of the unkown parameters specified at the end of section 2 proceeds in a number of steps.

1. Estimate disability risk directly from transitions between disability states

2. Estimate effect of disability on wages using wage data, controlling for selection

3. Estimate productivity risk from unexplained innovations to wages

4. Use indirect inference for the remaining parameters:

    - Estimate utility cost of disability, utility cost of participation, labour market frictions and the parameters of the disability insurance process
    - Use a range of auxilliary equations (coefficients from consumption regression, participation over the life-cycle, health status of DI recipients and the flows onto and off DI)

[11] We use a third-degree polinomial in $\ln F$ and control for a quartic in age, number of children, family size, dummies for white, education, region, year, and a quadratic in log before-tax family income.

[12] In future drafts we will use actual consumption, rather than imputed consumption data. Beginning in 1999, the PSID has added questions aimed at measuring spending on several categories of consumption (food, utilities, rent, mortgage payments, health care, child care, public and private transportations, and education). The PSID now covers about 70% of total expenditure as measured in the CEX. Li et al. (2007) show that each of the broad spending categories in the PSID aligns closely with the corresponding measures from the CEX. There are now four years of data available (1999, 2001, 2003, and 2005). Since we use disability status and disability insurance data also from an earlier period, we will need to make the assumption that the link between consumption, disability shocks, disability insurance and labor market participation is stable over time (so that it can be identified with the proposed strategy using only data from the more recent surveys).

4.1 Disability Risk

Disability risk is independent of any choices made by individuals in our model, and is also independent of productivity shocks. This means that the disability risk process can be identified sturcturally without indirect inference. By contrast, the same is not true for the variance of wage shocks which are identified using a selection correction that is based on a reduced form rather than on our structural model. We may include the wage risk parameters in the indirect inference estimation but we do not have to include the disability risk parameters.

4.2 The Wage Process

As said earlier, we model the wage process as being subject to general productivity shocks and shocks to the disability status (as well as the contribution of observable characteristics X_{it}):

$$\ln w_{it} = X_{it}'\alpha + \beta_1 \mathbf{1}\left\{L_{it}=1\right\} + \beta_2 \mathbf{1}\left\{L_{it}=2\right\} + \varepsilon_{it} + \omega_{it} \tag{7}$$

where ω_{it} is an i.i.d. measurement error, and

$$\varepsilon_{it} = \varepsilon_{it-1} + \eta_{it}$$

We make the assumption that the two shocks η_{it} and ω_{it} are independent and that disability shocks are orthogonal to the general stochastic component of individual productivity. Our goal is to identify the variance of the productivity shock σ_η^2 as well as β_1 and β_2.

The major complication for identifying the variance of the productivity shock arises from non-participation. Wages are not observed for non-participants. Moreover, non-participation depends on wages. Finally, non-participation may depend directly on disability shocks as well as the expectation that the individual will apply for *DI* in the subsequent period (which requires being unemployed in the current period). We observe neither these expectations, nor the decision to apply. Our approach is hence to write a reduced form model of participation:

$$\begin{aligned} P_{it}^* &= X_{it}'\gamma + \delta_1 \mathbf{1}\left\{L_{it}=1\right\} + \delta_2 \mathbf{1}\left\{L_{it}=2\right\} + \theta A_{it} + \pi_{it} \\ &= s_{it} + \pi_{it} \end{aligned} \tag{8}$$

where P_{it}^* is the utility from working, and we observe the indicator $P_{it} = \mathbf{1}\left\{P_{it}^* > 0\right\}$. Here A_{it} serves as an exclusion restriction: It affects the likelihood of observing an individual at work (through an income effect and through affecting the expectation that the individual will apply for *DI* in the subsequent period), but it does not affect the wage, conditional on X_{it} and L_{it}. The unobserved "taste for work" π_{it} is correlated with the permanent productivity component ε_{it}. We assume that

$$\begin{pmatrix} \varepsilon_{it} \\ \pi_{it} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\pi} \\ & 1 \end{pmatrix} \right)$$

The wage for labor market participants is thus:

$$\begin{aligned} E\left(\ln w_{it} | P_{it}^* > 0, X_{it}, L_{it}\right) &= X_{it}'\alpha + \beta_1 \mathbf{1}\left\{L_{it}=1\right\} + \beta_2 \mathbf{1}\left\{L_{it}=2\right\} + E\left(\varepsilon_{it} | P_{it}^* > 0, X_{it}, L_{it}\right) \\ &= X_{it}'\alpha + \beta_1 \mathbf{1}\left\{L_{it}=1\right\} + \beta_2 \mathbf{1}\left\{L_{it}=2\right\} + E\left(\varepsilon_{it} | \pi_{it} > -s_{it}\right) \\ &= X_{it}'\alpha + \beta_1 \mathbf{1}\left\{L_{it}=1\right\} + \beta_2 \mathbf{1}\left\{L_{it}=2\right\} + \sigma_{\varepsilon\pi}\lambda\left(s_{it}\right) \end{aligned}$$

assuming no selection on the measurement error. The Mills' ratio term $\lambda(s_{it}) = \frac{\phi(s_{it})}{\Phi(s_{it})}$, where $\phi(.)$ and $\Phi(.)$ denote the p.d.f. and c.d.f. of the standard normal distribution, respectively. Thus, we estimate

$$\ln w_{it} = X_{it}'\alpha + \beta_1 \mathbf{1}\{L_{it} = 1\} + \beta_2 \mathbf{1}\{L_{it} = 2\} + \sigma_{\varepsilon\pi}\lambda(s_{it}) + v_{it} \tag{9}$$

only on the sample of workers, and with $E(v_{it}|P_{it}^* > 0, X_{it}, L_{it}) = 0$. The resulting estimates of β_1 and β_2 should be interpreted as the estimates of the effect of work limitations on wage offers.

4.3 Productivity Risk

To identify the variance of productivity shocks, we define first the "adjusted" error term:

$$g_{it} = \Delta\left(\ln w_{it} - X_{it}'\alpha - \beta_1 \mathbf{1}\{L_{it} = 1\} - \beta_2 \mathbf{1}\{L_{it} = 2\}\right)$$

and then identify the variance of productivity shocks and the variance of measurement error using the following moment restrictions:

$$\begin{aligned} E(g_{it}|P_{it} = 1, P_{it-1} = 1) &= \rho_{\eta\pi}\sigma_\eta\lambda(s_{it}) & (10) \\ E\left(g_{it}^2|P_{it} = 1, P_{it-1} = 1\right) &= \sigma_\eta^2\left(1 - \rho_{\eta\pi}^2 s_{it}\lambda(s_{it})\right) + 2\sigma_\omega^2 & (11) \\ -E(g_{it}g_{it+1}|P_{it} = 1, P_{it-1} = 1) &= \sigma_\omega^2 & (12) \end{aligned}$$

(see Low, Meghir and Pistaferri, 2006). Here $\rho_{\eta\pi}$ denotes the correlation coefficient between η and π (which is not of direct interest).

This stage of the estimation proceeds is in three steps. In the first step we estimate a probit model for participation. This allows us to construct a consistent estimate of the inverse Mills' ratio. In the second step we regress log wages on X, dummies for work limitations, and the Mills ratio. We then construct the "adjusted" residuals, use them as they were the true "adjusted" error terms (MaCurdy, 1982), and estimate σ_η^2 and σ_ω^2 using (10)-(12). Standard errors are computed with the block bootstrap.

4.4 Preferences and Disability Insurance Parameters

Identification of the remaining structural parameters of interest $(\eta, \theta, \delta, F_{L=0}, F_{L=1}, F_{L=2})$ and the "policy" parameters $(P_{L=0}^{S,Y}, P_{L=1}^{S,Y}, P_{L=2}^{S,Y}, P_{L=0}^{S,O}, P_{L=1}^{S,O}, P_{L=2}^{S,O}$, and $P^{\text{Re}})$ will be achieved by Indirect Inference (see Gourieroux et al, 1993; Smith, 2006). Indirect inference is a simulation-based method that is used when the relevant theoretical moments have no analytical expressions. This is indeed the case for our complex theoretical model. The difference between indirect inference and other methods based on simulations (such as Simulated Method of Moments) is that indirect inference requires only the specification of an approximate model (known as auxiliary model). The auxiliary model is not necessarily the correct data generating process. However, the main idea behind indirect inference is that the parameters of the auxiliary model are related (through a so-called binding function) to the structural parameters of interest. The latter are estimated by minimizing the distance between the parameters of the auxiliary model estimated from the observed data and the parameters of the auxiliary model estimated from the simulated data.

We use the following Indirect Inference auxiliary equations, which overall give us 35 moments:

1. Regression of log consumption on work limitation, disability insurance, participation (and interactions);
2. Participation rates over the life-cycle, conditional on work limitation;
3. Stock of recipients of DI, conditional on disability status and age;
4. Flows onto and off DI, conditional on disability status and age.

The Indirect Inference statistical criterion is:

$$\hat{\phi} = \arg\min_{\phi} \left(\widehat{\alpha}^D - S^{-1} \sum_{s=1}^{S} \widehat{\alpha}^S \left(\phi \right) \right)' \Omega \left(\widehat{\alpha}^D - S^{-1} \sum_{s=1}^{S} \widehat{\alpha}^S \left(\phi \right) \right)$$

where $\widehat{\alpha}^D$ are the moments in the data, $\widehat{\alpha}^S (\phi)$ are the corresponding simulated moments (which are averaged over S simulations, $\alpha (\phi)$ is the binding function relating the auxiliary parameters to the structural parameters) for given parameter values ϕ, and Ω is the weighting matrix (diagonal of the covariance matrix from the data).

Standard errors of the structural parameters can be computed using the formula provided in Gourieroux et al. (1993), i.e.,

$$var \left(\widehat{\phi} \right) = \left(1 + \frac{1}{S} \right) \left(\frac{\partial \widehat{\alpha}^S (\phi)}{\partial \phi}' \Omega \frac{\partial \widehat{\alpha}^S (\phi)}{\partial \phi} \right)^{-1}$$

If $\dim (\alpha) > \dim (\phi)$, the model generates overidentifying restrictions that can be used to test the model. One can also test for local identification by computing the Jacobian matrix of the binding function and testing whether the matrix has full row rank. In what follows we discuss the mapping between structural and auxiliary parameters.

4.4.1 Moments: Consumption Regression

Disability is likely to have two separate effects on consumption: first, disability affects earnings and hence consumption through the budget constraint. The size of this effect will depend on the extent of insurance, both self-insurance and formal insurance mechansims, such as DI. The extent of insurance from DI obviously depends on being admitted onto the program, but conditional on receiving DI, the extent of insurance is greater for low income individuals because of the progressivity of the system through the AIME calculation.

The second possible effect of disability on consumption is through nonseparabilities in the utility function. For example, if being disabled reduces the marginal utility of consumption (eg through a loss of appetite) then consumption will fall on disability even if there is full insurance and marginal utility is smoothed over states of disability.

It is important to separate out these two effects. Stephens (2001) calculates the effect of the onset of disability on consumption, but does not distinguish whether the effect is through nonseparability or through the income loss directly.

Our method for separating out these two effects is to include the following regression as an auxilliary equation:

$$\begin{aligned} \ln c_{it} &= \alpha_0 + \alpha_1 L_{it}^1 + \alpha_2 L_{it}^1 DI_{it} + \alpha_3 L_{it}^2 + \alpha_4 L_{it}^2 DI_{it} \\ &\quad + \alpha_5 Y_{it}^P + \alpha_6 t + \alpha_7 t^2 + \alpha_8 A_{it} + \alpha_9 P_{it} + \upsilon_{it} \end{aligned}$$

The effect of a (moderate) disability on consumption is the combination of two effects: α_1 captures the full effect on consumption of individuals who are not insured. In the presence of full insurance ($DI = 1$), consumption associated with the moderate disability is increased by α_2 and so $(\alpha_1 + \alpha_2)$ captures the nonseparable part. The coefficients α_3 and α_4 correspond to the effects for a severe disability. We control for permanent income and age because we want to compare individuals facing the same level of insurance through the DI system. We control for unearned income to compare individuals with the same potential for self-insurance. The split between α_1 and α_2 is clear when insurance is full. More generally, if insurance is partial, then $(\alpha_1 + \alpha_2)$ captures both the non-separable part and the lack of full insurance for those receiving DI. However, the degree of partial insurance through DI depends on permanent income and age through the AIME formula. Indirect inference exploits this identification intuition without putting a structural interpretation on the values of α.

We can construct Y^P_{it} by using the information on individual wages available from entry into the PSID sample until the particular observation at age t.

Participation in the labour force can also provide insurance against disability shocks. In addition, participation has a direct effect on the marginal utility of consumption. We use α_9,combined with the average participation rates over the life-cycle, to capture this non-separable component and the fixed cost of work.

4.4.2 Moments: Participation over the Life-Cycle

We calculate participation rates by age and by disability status. This is equivalent to do the follwoing

$$p^L_{ia} = \sum_{x=1}^{X} \beta^L_x \mathbf{1}\left\{age_i \in x, L\right\} + \varepsilon_{ia}$$

where p_{ia} is an indicator for whether the person x denote the age bands and there are overall X age bands (say 25-34, 35-44, etc.). The moments we use are the β^L_x. We assume there is no heterogeneity among firms and wages are determined by individual productivity and disability status. We also use duration of unemployment by age of entry in the unemployment state and work limitation state at the age of entry.

These moments are related to fixed cost of participation with different disabilities, $F(L)$, the utility cost of participation,η,and the labor market frictions.

Frictions are identified by average labor market participation and unemployment duration over the life cycle. To see the intuition, consider a world in which there are no food stamps. Because people are born healthy and have no assets to finance consumption during unemployment, the decision not to work in the first period is infinitely costly in terms of utility. Hence, if we see people not working in the first period of life this must reflect lack of offers. More generally, labor market participation at young ages is informative about λ.Similarly, transitions out of work in the first periods of the life cycle are informative about the job destruction rate δ (because the reservation asset value is very high at young ages and nobody quits). Finally, the differences in participation across disability status groups is informative about the disability status-specific fixed costs of work (i.e., it is more costly to accommodate disabled workers than healthy workers in a work environment).

4.4.3 Moments: Stock of Recipients of Disability Insurance,

Stock of DI recipients (age < 45 and age ≥ 45) by work limitation status: $\mathrm{Fr}(DI_t = 1\,|L_{it})$

4.4.4 Moments: Flows onto and off Disability Insurance

To give an idea of what kind of variability we are exploiting, suppose we want to identify $P^{S,O}_{L=2} = P\left(DI_t = 1|DI^{App}_t = 1, L_t = 2, t > 45\right)$. What we observe in the data is

$$\begin{aligned} & P\left(DI_t = 1|\, DI_{t-1} = 0, L_t = 2, t > 45\right) \\ = \; & P^{S,O}_{L=2} \times P\left(DI^{App}_t = 1|, DI_{t-1} = 0, L_t = 2, t > 45\right) \end{aligned}$$

(because nobody gets DI without an application and only non-DI recipients apply for DI). It follows that:

$$\begin{aligned} P^{S,O}_{L=2} &= \frac{P\left(DI_t = 1|\, DI_{t-1} = 0, L_t = 2, t > 45\right)}{P\left(DI^{App}_t = 1|, DI_{t-1} = 0, L_t = 2, t > 45\right)} \\ &\geq P\left(DI_t = 1|\, DI_{t-1} = 0, L_t = 2, t > 45\right) \end{aligned}$$

Hence one can interpret the observable $P\left(DI_t = 1|\, DI_{t-1} = 0, L_t = 2, t > 45\right)$ (the flow into DI by older workers who are severely work limitated) as a lower bound for the unobserved theoretical parameter $P^{S,O}_{L=2}$ (the fraction of older workers with a severel work limitation who get admitted into the DI program following an application). The tightness of the "bound" depends on how close $P\left(DI^{App}_t = 1|, DI_{t-1} = 0, L_t = 2, t > 45\right)$ is to 1. If all older workers with a severel work limitation not on DI at time $t - 1$ were to apply for DI, the two parameters would coincide. The theoretical model acts as a filter between the two, because applying to DI has important opportunity costs (that is, changes in the parameters of the model imply changes in $P\left(DI^{App}_t = 1|, DI_{t-1} = 0, L_t = 2, t > 45\right)$).

The moments we use are the flows onto DI, given by the fraction observed to start receiving DI: $Fr\left(DI_t = 1|\, DI_{t-1} = 0, L_t\right)$; and the flows off DI: $Fr\left(DI_t = 0|\, DI_{t-1} = 1, L_t\right)$ (either voluntarily or involuntarily). These moments (together with the stock of recipients of DI) are related to the structural parameters $P^{S,Y}_{L=0}, P^{S,Y}_{L=1}, P^{S,Y}_{L=2}, P^{S,O}_{L=0}, P^{S,O}_{L=1}, P^{S,O}_{L=2}, P^{\mathrm{Re}}$. For examples, $Fr\left(DI_t = 0|\, DI_{t-1} = 1, L_t\right)$ is clearly connected to the probability of reassessment P^{Re}. The connection is not one-to-one because our model suggests there are conditions under which a voluntary exit from DI is optimal (i.e., an extremely positive shock to individual productivity reflected in a higher wage offer if employed, perhaps linked with a medical recovery).

5 Results

5.1 Disability Risk

Figures 6-8 plot $\Pr\left(L_{it} = j|L_{it-1} = k\right)$ for $k = \{0, 1, 2\}$, respectively. These are transition probabilities that are informative about the "disability risk". For example, $\Pr\left(L_{it} = 2|L_{it-1} = 0\right)$ is the probability that an individual with no work limitations is hit by a shock that places him in

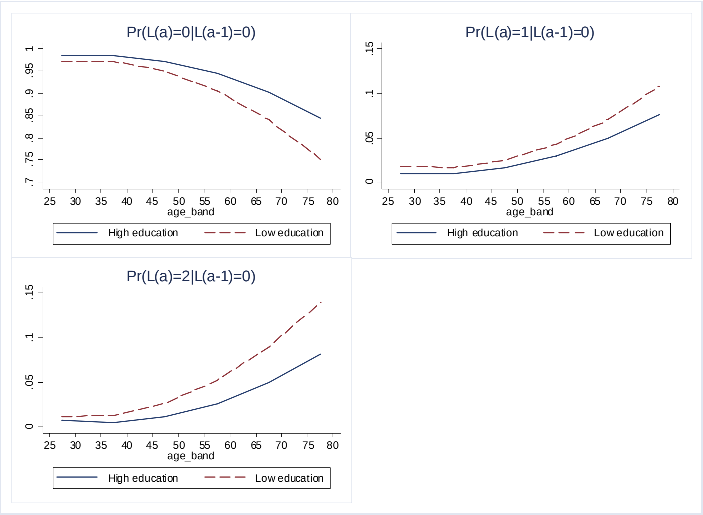

Figure 6: (Smoothed) Markov transition probabilities $\Pr(L_{ia} = j | L_{ia-1} = 0)$, by education.

the severe work limitations category. Whether this is a persistent or temporary transition can be answered by the value of $\Pr(L_{it} = 2 | L_{it-1} = 2)$.

In Figure 6 we start by plotting $\Pr(L_{it} = j | L_{it-1} = 0)$, i.e., the transition probabilities from the state of "no work limitations". To avoid cluttering, we have taken 5-age averages (25-29, 30-34, etc.). We also plot the predicted value of a regression on a quadratic in age. The probability of staying without work limitations declines over the working part of the life cycle from 0.98 to about 0.94. The decline is equally absorbed by increasing probabilities of transiting in moderate and severe work limitations. Figure 7 plots $\Pr(L_{it} = j | L_{it-1} = 1)$, i.e., the probability of transiting from a state of moderate work limitations. The probability of getting better declines over the life-cycle, while the probability of getting worse increases, especially after age 45. The probability of remaining with moderate work limitations has a U-shape. Finally, in Figure 8 we plot $\Pr(L_{it} = j | L_{it-1} = 2)$,where the transition is from the state of severe work limitations. Both the probability of a slight recovery ($\Pr(L_{it} = 1 | L_{it-1} = 2)$) and the probability of a strong recovery ($\Pr(L_{it} = 0 | L_{it-1} = 2)$) decline with age. In other words, persistence in the severe work limitations state increases with age. The low educated face worse health risk than the high educated group, with higher probabilities of bad shocks occuring and a lower probability of recovering.

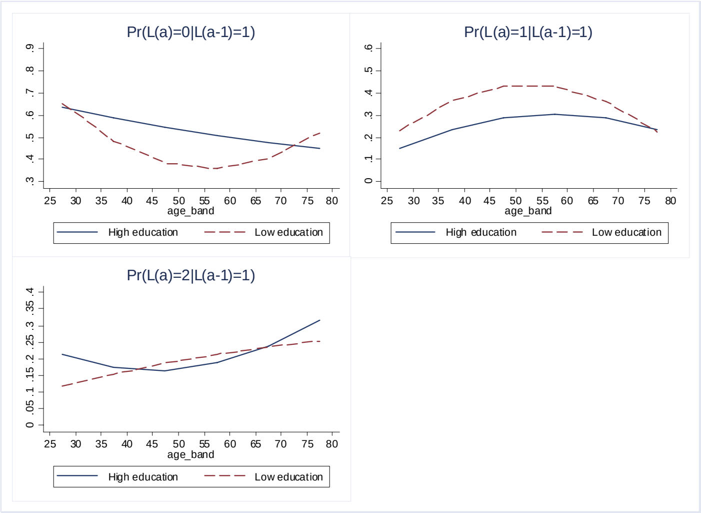

Figure 7: (Smoothed) Markov transition probabilities $\Pr\left(L_{ia} = j|L_{ia-1} = 1\right)$, by education.

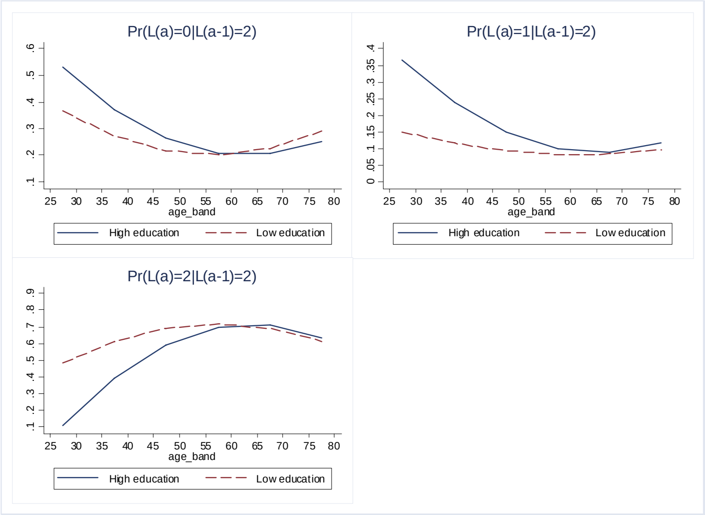

Figure 8: (Smoothed) Markov transition probabilities $\Pr\left(L_{ia} = j|L_{ia-1} = 2\right)$, by education.

5.2 Wage Process

Table 1: Participation Probit

Variable	Low education	High education
L_{it}= 1	-0.182 (0.026)	-0.076 (0.025)
L_{it}= 2	-0.588 (0.026)	-0.358 (0.049)
Age	0.008 (0.002)	0.004 (0.001)
$\frac{Age^2}{100}$	-0.011 (0.002)	-0.006 (0.001)
White	0.055 (0.006)	-0.007 (0.004)
Married	0.054 (0.009)	0.017 (0.005)
$\frac{\text{Unearned income}}{1000}$	-0.006 (0.0004)	-0.001 (0.0001)
N	8836	7103

In Table 1 we report the results of estimating a probit regression for participation separately for low education individuals (those with a HighSchool diploma or less) and high education individuals (with at least some college education). Participation is monotonically decreasing in the degree of work limitations. We report marginal effects. Thus, the interpretation is that among the low educated, the probability of working declines by 0.18 units at the onset of moderate work limitations, and by 0.59 units at the onset of severe work limitations. The figures among the high educated are slightly smaller, 0.08 and 0.36, respectively. As for our exclusion restriction, its sign is correct (higher unearned income should increase the opportunity cost of work), and the effect is statistically significant in both groups. The other effects have signs that are consistent with previous evidence.

In Table 3 we report estimates of the log wage process with and without correcting for endogenous selection into work. This makes a substantial difference. For both groups the wage loss associated with the onset of work limitations is higher when selection is being taken into account. The sign of the Mills' ratio suggests positive selection on unobservables (i.e., people with bad realization of their permanent component quit into unemployment - or are laid-off), and it is statistically significant. The reason for the selection bias is simple. From Table 2, an increase in work limitations pushes some people out of work. Those who leave work tend to be those with low unobserved propensity to work (i.e., low π_{it}), which also tend to be individuals with low unobserved permanent income (i.e., low ε_{it}). Hence, if one ignores selection, the wage loss associated with an increase in work limitations appears attenuated by the fact that, among those with work limitations, those who remain at work are higher-than-average permanent income people. Once selection is taken into account, the full loss of disability is revealed.

Table 2: The log wage equation

Variable	Low education		High education	
$\{L_{it}=1\}$	-0.266 (0.036)	-0.335 (0.038)	-0.200 (0.058)	-0.296 (0.060)
$\{L_{it}=2\}$	-0.332 (0.050)	-0.564 (0.074)	-0.310 (0.070)	-0.637 (0.098)
Age	0.059 (0.004)	0.067 (0.005)	0.088 (0.006)	0.106 (0.007)
$\frac{Age^2}{100}$	-0.062 (0.005)	-0.072 (0.006)	-0.085 (0.008)	-0.109 (0.009)
White	0.209 (0.011)	0.226 (0.012)	0.229 (0.013)	0.245 (0.013)
Married	0.130 (0.015)	0.150 (0.015)	0.127 (0.018)	0.158 (0.018)
Mills ratio		0.208 (0.049)		0.577 (0.117)
N	7964	7964	6882	6882

5.3 Productivity Risk

We use the residuals of the wage equation to estimate the variance of permanent productivity shocks as well as the variance of transitory shocks, allowing for endogenous selection into work (expressions (10)-(12)). The results are in Table 3. We find that the variance of permanent shocks is slightly higher among the low educated. For the variance of measurement error (or transitory shocks) the ranking is reversed.

Table 3: The variances of the productivity shocks

Variable	Low education	High education
	(1)	(2)
Permanent shock	0.024 (0.0068)	0.036 (0.0071)
Measurement error (Transitory)	0.030 (0.0031)	0.022 (0.0032)

5.4 Estimates from Indirect Inference

Here we report the estimates we obtain using Indirect Inference. First, we set some parameters to realistic values (Table 4):

Table 4: Exogenous Parameters

	Value
γ	1.5
R	0.016 (annual)
β	0.025 (annual)
T	200 (40 years)
R	40 (10 years)
λ	0.73

In future drafts we will estimate λ using unemployment duration data. In this draft we are using only average participation rates by age and work limitation status, and so λ is hard to pin down.

Next, we present results from estimating the auxiliary log consumption equation (using imputed data), see Table 5. We obtain a good match between data and simulations. The signs and in most cases even the magnitude of the coefficients are similar. These numbers are not intrinsically interesting, however. It is their link with structural parameters that it is more interesting for our purposes.

Table 5: The Log Consumption Equation

Variable	Low Education	
	Data	Simulations
$\{L_{it} = 1\}$	-0.124	-0.103
$\{L_{it} = 1\}\, DI$	0.030*	0.299
$\{L_{it} = 2\}$	-0.182	-0.190
$\{L_{it} = 2\}\, DI$	0.165	0.414
Employed	0.183	0.370

Controls: Age, Age^2,Unearned income, Permanent income
A * denotes a statistically insignificant estimate.

Table 6 shows participation over the life cycle for people in different work limitation categories. Our simulations match quite well participation of the non-disabled, less well that of the disabled. In particular, young individuals (in the simulations) do not work when severely limited, whereas in the data, more of those individuals work. Old individuals (in the simulations) work when severely limited rather than applying for DI compared to the data. These discrepancies will be studied and hopefully understood and corrected in future drafts.[13]

Table 6: Labor Market Participation by Disability Status

Age	No limitation		Moderate limitation		Severe limitation	
	Data	Simulations	Data	Simulations	Data	Simulations
22-31	0.92	0.94	0.65	0.67	0.38	0.13
32-41	0.93	0.93	0.75	0.69	0.28	0.31
42-51	0.95	0.91	0.72	0.72	0.23	0.32
52-61	0.84	0.87	0.45	0.66	0.10	0.26

[13] For example, the model underpredicts employment of young workers with severe disability. One way to reduce this degree of underprediction is to reduce the wage penalty associated with being disabled by imposing a minimum wage floor. In the current model wages may drop to such low levels in response to disability that very few workers with severe disability may decide to work. In reality, the fall of wages is limited by, say, the presence of a minimum wage. Relatedly, we will check whether the ratio of food stamps to minimum wage earnings in the simulations is consistent with that found in the data. If this ratio is too high in the model, it may represent an incentive not to work given severe disability.

The last piece of evidence comes from matching DI recipiency moments. In Table 7 there are three sets of moments: the Stock of DI recipient, the Flow into DI, and the Flow off DI, all by disability status. There are cases in which the fit of the model is excellent (i.e., the stock of DI recipients, and the flows off DI at young ages). In other cases, the fit of the model needs to be improved (i.e., flows onto DI at old ages). The fact that the fit of the model needs to be improved is also confirmed by the fact that the overidentifying restrictions rare rejected $\left(\chi^2_{23} = 69\right)$.

Table 7: Moments Associated with the Disability Insurance Process

		23-45		46-62	
		Data	Simul	Data	Simul
Stock of DI Recipients	$L = 0$	0.005	0.002	0.025	0.013
	$L = 1$	0.050	0.038	0.074	0.091
	$L = 2$	0.25	0.27	0.51	0.31
Flows onto DI	$L = 0$	0.001*	0.0002	0.0072	0.0009
$P\left(DI_t = 1 \mid DI_{t-1} = 0, L_t\right)$	$L = 1$	0.006*	0.0011	0.029	0.016
	$L = 2$	0.11	0.14	0.20	0.11
Flows off DI	$L = 0$	0.31*	0.78	0.17*	0.48
$P\left(DI_t = 0 \mid DI_{t-1} = 1, L_t\right)$	$L = 1$	0.33*	0.31	0.29*	0.22
	$L = 2$	0.15	0.11	0.14	0.04

Note: A * denotes a statistically insignificant estimate

In Table 8 we report the Indirect Inference estimates obtained by minimizing the distance between the moments computed from the data (i.e., those reported in Tables 4, 5, and 6), and the equivalent moments computed from the simulated model. We estimate that disability induces about a 2% loss of utility in terms of consumption. Participation induces a 32% loss. The fixed costs are reported as the fraction of average offered wage income at age 22. They rise with the degree of disability. We estimate that a job is destroyed on average every 26 quarter and that a job offer is received by the unemployed every 1.4 quarters. The probability of success of DI application increases with age and disability status. Each DI recipients faces a 10% probability of being re-assessed.

Table 8: Estimated Parameters

Frictions and Preferences			Disability Insurance Program	
Paramter		Estimate	Parameter	Estimate
θ	Cost of disability	-0.017	$P^{S,Y}_{L=0}$	0.065
η	Cost of part.	-0.32	$P^{S,Y}_{L=0}$	0.140
δ	Job destruction	0.038	$P^{S,Y}_{L=1}$	0.075
$F_{L=0}$	Fixed cost	0.24	$P^{S,O}_{L=1}$	0.260
$F_{L=1}$	Fixed cost	0.42	$P^{S,Y}_{L=2}$	0.465
$F_{L=2}$	Fixed cost	0.74	$P^{S,O}_{L=2}$	0.925
			P^{Re}	0.092

Note: Fixed costs are reported as the fraction of average offered wage income at age 23. All parameters significant (using asymptotic standard errors, not correcting for first stage estimates)

6 Implications

Our theoretical framework and structural estimates of the model can be used to study the implications of the existing DI program, as well as to evaluate the welfare effects of modifying the features of the current program.

6.1 Success of the DI Screening Process

One important issue is to evaluate the success rate of the current DI Screening Process. Let's start from the Award rate, $\Pr(DI = 1|DI^{App} = 1)$. We estimate this rate (using our structural model and estimated parameters) to be 0.53. This contrasts quite well with the reduced form estimates (0.45) obtained by Bound and Burkhauser (1999) and others using data on DI application and DI receipt from the HRS.

Given that the true disability status of an applicant is private information, SS evaluators are bound to commit two types of errors: Admitting into the DI program undeserved applicants and rejecting those who are truly disabled. How large are the probabilities associated with these errors? Consider first the extent of false positives (the proportion of healthy individuals who apply receiving DI). We estimate the following probabilities:

$$\begin{aligned} Pr(DI &= 1|L=0, DI^{App}=1, age \geq 45) = 0.14 \\ Pr(DI &= 1|L=1, DI^{App}=1, age \geq 45) = 0.26 \end{aligned}$$

What about the Award Error? This is $Pr(L = \{0,1\}|DI = 1, DI^{App} = 1) = 0.10$. In the literature, we have found reduced form estimates that are fairly similar, 0.18 in Benitez-Silva et al. (1999), 0.22 in Benitez-Silva, Bushinsky, Rust (2006), and 0.19 in Nagi (1969).

Consider next the probability of false negatives (i.e., the proportion of severely disabled who apply and do not receive DI). We estimate:

$$\begin{aligned} Pr(DI &= 0|L=2, DI^{App}=1, age \geq 45) = 0.07 \\ Pr(DI &= 0|L=2, DI^{App}=1, age < 45) = 0.53 \end{aligned}$$

The Rejection Error is $Pr(L=2|DI=0, DI^{App}=1) = 0.43$. Contrast this with Benitez-Silva et al. (1999), 0.50, Benitez-Silva, Bushinsky, Rust (2006), 0.58, and Nagi (1969), 0.48. These comparisons confirm that our structural model is capable of replicating quite well reduced form estimates obtained using direct information on the application and award process.

6.2 Changing Parameters of the DI Process

The most important use of our model is the ability to measure the welfare effects of changing the main parameters of the DI programs. Consider making the program "stricter". In one form or another, this suggestion has been advanced as one possible solution to the "moral hazard" problem. To tackle this issue, one needs to define first a measure of strictness of the program. Suppose that Social Security DI evaluators decide whether to award DI as a function of a signal about the applicant's disability status:

$$S_{it} = \alpha_{t,L} + \xi_{it}$$

The mean of the signal ($\alpha_{t,L}$) varies by age (for simplicity, for two age groups defined by age<45 and age≥45), and by work limitation status L. ξ is a normally distributed error with variance σ_ξ^2. Assume that the Social Security DI evaluators decide to award DI if $S_{it} > \overline{S}$. The parameter can be interpreted as a measure of strictness of the DI program (ceteris paribus, an increase in reduces the proportion of people admitted into the program). Note that this framework connects the estimated structural probabilities described below ($P_L^{S,t}$) with the parameters $\overline{S}$, $\alpha_{t,L}$, and σ_ξ^2, i.e. through

$$\Phi\left(\frac{\overline{S} - \alpha_{t,L}}{\sigma_\xi}\right) = 1 - P_L^{S,t}$$

where $\Phi(.)$ is the c.d.f. of the standard normal. Using the 6 probabilities of acceptance (by type and age) from the estimation and using the normalizations $\alpha_{O,L=2} = 1, \alpha_{O,L=0} = 0$, one can solve to find estimates of the threshold $\overline{S}$, $\alpha_{t,L}$, and σ_ξ^2 (for $t = \{$"Y" or age<45, and "O" or age≥ 45$\}$ and $L = \{0, 1, 2\}$). Figure 9 illustrates the extent of errors under the current DI program. The area on the left of under the blue curve (the one labeled $f(S|L=2, t \geq 45)$) measures the probability of rejecting a deserving DI applicant. The areas on the right of under the black and red curves (the ones labeled $f(S|L=0, t \geq 45)$ and $f(S|L=1, t \geq 45)$, respectively) measure the probability of accepting into the DI program an undeserving DI applicant. Increasing the strictness of the test (increasing) reduces the probability of type II error (reduces the extent of the moral hazard problem), but also increases the probability of type I error (reduces the extent of insurance provided by the program). This is a classical conundrum in hypothesis testing. Since this policy has both benefits and costs, one important element of our project is to use our model to determine whether an increase in the strictness of the test would be welfare-improving or welfare-worsening.

We consider the following strategy. We consider the effect of changes in the acceptance threshold . We hold the government's budget constant, which is achieved by adjusting the proportional payroll tax (this is done iteratively because labor supply changes as a consequence). We calculate expected utility and the extent of moral hazard (false applications) for different values of the acceptance threshold $\overline{S}$.[14]

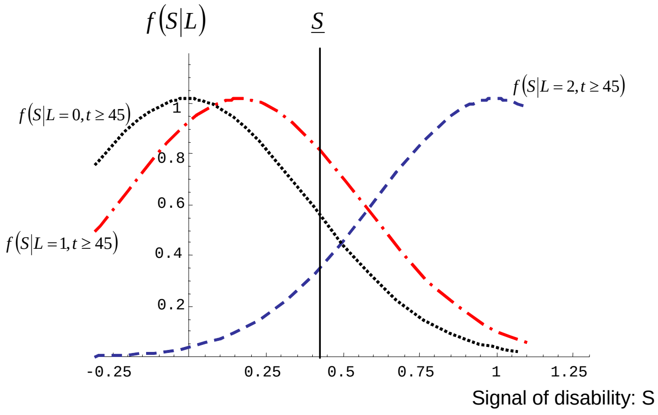

Figure 9:

Figure 10 reports the results of this experiment. We find that the "optimal" acceptance threshold lies on the right of the estimated (actual) threshold, i.e., increasing the strictness of the test is welfare-improving (note that, absent a theoretical framework, nothing could be said about whether the optimal threshold is higher or lower than the estimated one). In the optimal scenario (obtained by maximization of expected utility *beyond the veil of ignorance*, i.e., before individuals discover their types etc.), the acceptance threshold is about 50% higher than the estimated one. Hence, we find that it is welfare enhancing to make the medical test stricter to reduce false positives (and moral hazard), despite the worsening in the degree of insurance provided.

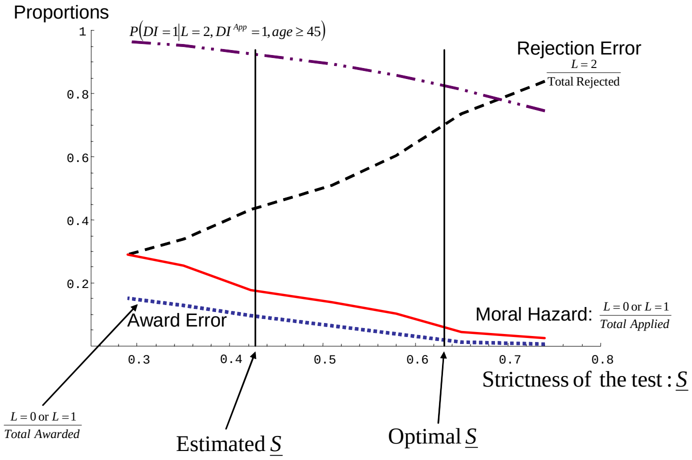

Figure 10:

We have also considered changing other parameters of the DI program (results are preliminary):

[14] An alternative is, of course, to invest in technologies that increase the degree of information about individuals' true disability status (i.e., reducing σ_{ξ}^2). Unfortunately, this policy is typically very expensive.

1. Reducing the generosity of payments. This reduces moral hazard and award errors. People are willing to pay to reduce generosity (and obtain a tax cut as a consequence);
2. Increasing the reassessment rate. This induces a small reduction in moral hazard and award errors. Also in this case, people appear willing to pay to move into this new scenario;
3. Reducing food stamps benefits. This worsens moral hazard because there is a fall in applications coming from L=2 people. People require compensation for reduced food stamps (despite the lower payroll tax).

These preliminary results should be taken with caution for a number of reasons. First, we are working on extending the data to the 1987-2005 period where we have access to better data on consumption (and perhaps a new "steady state" following the mid-1980 policy interventions that liberalized access into DI). Second, we are estending the model to include SSI and a better characterization of the consumption floor program. Third, we have a made a number of simplifications (no health investments, for example) whose effect needs to be assessed. Finally, the fit of the model (as described above) needs to be improved.

7 Conclusions

- Extent of disability risk
  - work limitations large impact on wage level (50% lower wage)
    and on participation decision
  - disability risk accounts for a small part of productivity risk
- Probabilities of acceptance onto DI program:
  - Those with severe limitations suffer low rejections
  - False positives (awards to the healthy) more problematic
- Use life-cycle model to explore trade-off between insurance and false applications
- Policy changes:
  - Welfare increasing to make the medical test more strict to reduce false positives (and moral hazard), despite worsening in insurance provided
- Among severely work limited, young individuals do not work enough (in the simulations), old individuals work too much
- What is the right specification of utility?
- Can we understand recent changes in DI enrollment through the model estimates? What has been happening to “strictness” of the tests?
- Investment in health: alternative margin for moral hazard

References

[1] Adda, J., Banks, J. and H-M von Gaudecker (2007) "The impact of income shocks on health: evidence from cohort data" Institute for Fiscal Studies Working Paper 07/05

[2] Autor, David and Mark Duggan (2006) "The growth in the social security disability rolls: a fiscal crisis unfolding" NBER Working Paper 12436

[3] Banks, James, Richard Blundell and Sarah Tanner (1998), "Is There a Retirement-Savings Puzzle?", American Economic Review, 88 (4), 769-788.

[4] Benitez-Silva, Hugo, Moshe Buchinsky, Hiu Man Chan, Sofia Cheidvasser, and John Rust (2000), "How Large Is the Bias in Self-Reported Disability?", Working Paper no. 7526 (February), NBER, Cambridge, MA.

[5] Benitez-Silva, Hugo, Moshe Buchinsky and John Rust (2004), "How large are the classification errors in the social security disability award process?" NBER Working Paper 10219

[6] Bernheim, B. Douglas, Jonathan Skinner and Steven Weinberg (2001), "What Accounts for the Variation in Retirement Wealth among U.S. Households?", American Economic Review, 91 (4), pp. 832-857.

[7] Bound, John and Richard V Burkhauser (1999) "Economic analysis of transfer programs targeted on people with disabilities" in: Handbook of Labor Economics, Volume 3, edited by: O. Ashenfelter and D. Card

[8] Bound, J., Cullen, J. B., Nichols, A. and L. Schmidt (2004) "The welfare implications of increasing disability insurance benefit generosity" Journal of Public Economics 88:2487-2514

[9] Browning, Martin and Thomas F. Crossley (2001), "Unemployment Benet Levels and Consumption Changes", Journal of Public Economics, 80(1), 1-23.

[10] Chen, S. and W. van der Klaauw (2005) "The work disincentive effects of the disability insurance program in the 1990s"

[11] Cochrane, John (1991), "A Simple Test of Consumption Insurance", Journal of Political Economy 99:5, 957-976.

[12] Cunha and Heckman (2006), "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation").

[13] Gruber, Jonathan (1997), "The Consumption Smoothing Benefits of Unemployment Insurance", American Economic Review, 87(1), 192-205.

[14] Golosov, Mikhail and Aleh Tsyvinski (2006), "Designing Optimal Disability Insurance: A Case for Asset Testing", Journal of Political Economy, volume 114, 257-279.

[15] Hurd, Michael D., and Susann Rohwedder (2006), "Some Answers to The Retirement-Consumption Puzzle", Rand Working paper WR-342.

[16] Jones, Andrew and Owen O'Donnell (1995) "Equivalence scales and the costs of disability" Journal of Public Economics 56:273-289

[17] Kreider, B. and J. Pepper (2001) "Inferring disability status from corrupt data"

[18] Meyer, Bruce D. and Wallace K.C. Mok (2007) "Disability, earnings, income and consumption" University of Chicago, mimeo

[19] Ruhm, Christopher J. (2001) "Economic expansions are unhealthy: evidence from microdata" NBER Working Paper 8447

[20] Smith, J. (2004) "Unravelling the SES health connection" Institute for Fiscal Studies Working Paper 04/02

[21] van der Klaauw, W. and K. Wolpin (2006) "Social Security and the retirement and savings behavior of low income households"

[22] Viscusi, W.Kip and William N. Evans (1990) "Utility functions that depend on health status: estimates and economic implications" American Economic Review 80(3):353-374